20	40	60
MGMSKSHSFFGYPLSIFFIV	VNEFCERFSYYGMRAILILY	FTNFISWDDNLSTAIYHTFV
80	100	120
ALCYLTPILGALIADSWLGK	FKTIVSLSIVYTIGQAVTSV	SSINDLTDHNHDGTPDSLPV
140	160	180
HVVLSLIGLALIALGTGGIK	PCVSAFGGDQFEEGQEKQRN	RFFSIFYLAINAGSLLSTII
200	220	240
TPMLRVQQCGIHSKQACYPL	AFGVPAALMAVALIVFVLGS	GMYKKFKPQGNIMGKVAKCI
260	280	300
GFAIKNRFRHRSKAFPKREH	WLDWAKEKYDERLISQIKMV	TRVMFLYIPLPMFWALFDQQ
320	340	360
GSRWTLQATTMSGKIGALEI	QPDQMQTVNAILIVIMVPIF	DAVLYPLIAKCGFNFTSLKK
380	400	420
MAVGMVLASMAFVVAAIVQV	EIDKTLPVFPKGNEVQIKVL	NIGNNTMNISLPGEMVTLGP
440	460	480
MSQTNAFMTFDVNKLTRINI	SSPGSPVTAVTDDFKQGQRH	TLLVWAPNHYQVVKDGLNQK
500	520	540
PEKGENGIRFVNTFNELITI	TMSGKVYANISSYNASTYQF	FPSGIKGFTISSTEIPPQCQ
560	580	600
PNFNTFYLEFGSAYTYIVQR	KNDSCPEVKVFEDISANTVN	MALQIPQYFLLTCGEVVFSV
620	640	660
TGLEFSYSQAPSNMKSVLQA	GWLLTVAVGNIIVLIVAGAG	QFSKQWAEYILFAALLLVVC
680	700	708
VIFAIMARFYTYINPAEIEA	QFDEDEKKNRLEKSNPYFMS	GANSQKQM

Fig. 1

```
   1 gaattccgtc tcgaccactg aatggaagaa aaggactttt aaccaccatt ttgtgactta
  61 cagaaaggaa tttgaataaa gaaaactatg atacttcagg cccatcttca ctccctgtgt
                                  M  I  L  Q  A  H  L  H  S  L  C
 121 cttcttatgc tttatttggc aactggatat ggccaagagg ggaagtttag tggacccctg
       L  L  M  L  Y  L  A  T  G  Y  G  Q  E  G  K  F  S  G  P  L
 181 aaacccatga cattttctat ttatgaaggc caagaaccga gtcaaattat attccagttt
       K  P  M  T  F  S  I  Y  E  G  Q  E  P  S  Q  I  I  F  Q  F
 241 aaggccaatc ctcctgctgt gacttttgaa ctaactgggg agacagacaa catatttgtg
       K  A  N  P  P  A  V  T  F  E  L  T  G  E  T  D  N  I  F  V
 301 atagaacggg agggacttct gtattacaac agagccttgg acagggaaac aagatctact
       I  E  R  E  G  L  L  Y  Y  N  R  A  L  D  R  E  T  R  S  T
 361 cacaatctcc aggttgcagc cctggacgct aatggaatta tagtggaggg tccagtccct
       H  N  L  Q  V  A  A  L  D  A  N  G  I  I  V  E  G  P  V  P
 421 atcaccatag aagtgaagga catcaacgac aatcgaccca cgtttctcca gtcaaagtac
       I  T  I  E  V  K  D  I  N  D  N  R  P  T  F  L  Q  S  K  Y
 481 gaaggctcag taaggcagaa ctctcgccca ggaaagccct tcttgtatgt caatgccaca
       E  G  S  V  R  Q  N  S  R  P  G  K  P  F  L  Y  V  N  A  T
 541 gacctggatg atccggccac tcccaatggc cagctttatt accagattgt catccagctt
       D  L  D  D  P  A  T  P  N  G  Q  L  Y  Y  Q  I  V  I  Q  L
 601 cccatgatca acaatgtcat gtactttcag atcaacaaca aaacgggagc catctctctt
       P  M  I  N  N  V  M  Y  F  Q  I  N  N  K  T  G  A  I  S  L
 661 acccgagagg gatctcagga attgaatcct gctaagaatc cttcctataa tctggtgatc
       T  R  E  G  S  Q  E  L  N  P  A  K  N  P  S  Y  N  L  V  I
 721 tcagtgaagg acatgggagg ccagagtgag aattccttca gtgataccac atctgtggat
       S  V  K  D  M  G  G  Q  S  E  N  S  F  S  D  T  T  S  V  D
 781 atcatagtga cagagaatat ttggaaagca ccaaaacctg tggagatggt ggaaaactca
       I  I  V  T  E  N  I  W  K  A  P  K  P  V  E  M  V  E  N  S
 841 actgatcctc accccatcaa aatcactcag gtgcggtgga atgatcccgg tgcacaatat
       T  D  P  H  P  I  K  I  T  Q  V  R  W  N  D  P  G  A  Q  Y
 901 tccttagttg acaaagagaa gctgccaaga ttcccatttt caattgacca ggaaggagat
       S  L  V  D  K  E  K  L  P  R  F  P  F  S  I  D  Q  E  G  D
 961 atttacgtga ctcagccctt ggaccgagaa gaaaaggatg catatgtttt ttatgcagtt
       I  Y  V  T  Q  P  L  D  R  E  E  K  D  A  Y  V  F  Y  A  V
1021 gcaaaggatg agtacggaaa accactttca tatccgctgg aaattcatgt aaaagttaaa
       A  K  D  E  Y  G  K  P  L  S  Y  P  L  E  I  H  V  K  V  K
1081 gatattaatg ataatccacc tacatgtccg tcaccagtaa ccgtatttga ggtccaggag
       D  I  N  D  N  P  P  T  C  P  S  P  V  T  V  F  E  V  Q  E
1141 aatgaacgac tgggtaacag tatcgggacc cttactgcac atgacaggga tgaagaaaat
       N  E  R  L  G  N  S  I  G  T  L  T  A  H  D  R  D  E  E  N
1201 actgccaaca gttttctaaa ctacaggatt gtggagcaaa ctcccaaact tcccatggat
       T  A  N  S  F  L  N  Y  R  I  V  E  Q  T  P  K  L  P  M  D
```

Fig. 2A

```
1261 ggactcttcc taatccaaac ctatgctgga atgttacagt tagctaaaca gtccttgaag
      G  L  F   L  I  Q   T  Y  A   G  M  L   Q  L  A   K  Q  S  L  K
1321 aagcaagata ctcctcagta caacttaacg atagaggtgt ctgacaaaga tttcaagacc
      K  Q  D   T  P  Q   Y  N  L   T  I  E   V  S  D   K  D  F  K  T
1381 ctttgttttg tgcaaatcaa cgttattgat atcaatgatc agatccccat ctttgaaaaa
      L  C  F   V  Q  I   N  V  I   D  I  N   D  Q  I   P  I  F  E  K
1441 tcagattatg gaaacctgac tcttgctgaa gacacaaaca ttgggtccac catcttaacc
      S  D  Y   G  N  L   T  L  A   E  D  T   N  I  G   S  T  I  L  T
1501 atccaggcca ctgatgctga tgagccattt actgggagtt ctaaaattct gtatcatatc
      I  Q  A   T  D  A   D  E  P   F  T  G   S  S  K   I  L  Y  H  I
1561 ataaagggag acagtgaggg acgcctgggg gttgacacag atccccatac caacaccgga
      I  K  G   D  S  E   G  R  L   G  V  D   T  D  P   H  T  N  T  G
1621 tatgtcataa ttaaaaagcc tcttgatttt gaaacagcag ctgtttccaa cattgtgttc
      Y  V  I   I  K  K   P  L  D   F  E  T   A  A  V   S  N  I  V  F
1681 aaagcagaaa atcctgagcc tctagtgttt ggtgtgaagt acaatgcaag ttcttttgcc
      K  A  E   N  P  E   P  L  V   F  G  V   K  Y  N   A  S  S  F  A
1741 aagttcacgc ttattgtgac agatgtgaat gaagcacctc aattttccca acacgtattc
      K  F  T   L  I  V   T  D  V   N  E  A   P  Q  F   S  Q  H  V  F
1801 caagcgaaag tcagtgagga tgtagctata ggcactaaag tgggcaatgt gactgccaag
      Q  A  K   V  S  E   D  V  A   I  G  T   K  V  G   N  V  T  A  K
1861 gatccagaag gtctggacat aagctattca ctgaggggag acacaagagg ttggcttaaa
      D  P  E   G  L  D   I  S  Y   S  L  R   G  D  T   R  G  W  L  K
1921 attgaccacg tgactggtga gatctttagt gtggctccat tggacagaga agccggaagt
      I  D  H   V  T  G   E  I  F   S  V  A   P  L  D   R  E  A  G  S
1981 ccatatcggg tacaagtggt ggccacagaa gtaggggggt cttccttaag ctctgtgtca
      P  Y  R   V  Q  V   V  A  T   E  V  G   G  S  S   L  S  S  V  S
2041 gagttccacc tgatccttat ggatgtgaat gacaaccctc ccaggctagc caaggactac
      E  F  H   L  I  L   M  D  V   N  D  N   P  P  R   L  A  K  D  Y
2101 acgggcttgt tcttctgcca tcccctcagt gcacctggaa gtctcatttt cgaggctact
      T  G  L   F  F  C   H  P  L   S  A  P   G  S  L   I  F  E  A  T
2161 gatgatgatc agcacttatt tcggggtccc catttacat tttccctcgg cagtggaagc
      D  D  D   Q  H  L   F  R  G   P  H  F   T  F  S   L  G  S  G  S
2221 ttacaaaacg actgggaagt ttccaaaatc aatggtactc atgcccgact gtctaccagg
      L  Q  N   D  W  E   V  S  K   I  N  G   T  H  A   R  L  S  T  R
2281 cacacagact ttgaggagag ggcgtatgtc gtcttgatcc gcatcaatga tgggggtcgg
      H  T  D   F  E  E   R  A  Y   V  V  L   I  R  I   N  D  G  G  R
2341 ccacccttgg aaggcattgt ttctttacca gttacattct gcagttgtgt ggaaggaagt
      P  P  L   E  G  I   V  S  L   P  V  T   F  C  S   C  V  E  G  S
2401 tgtttccggc cagcaggtca ccagactggg atacccactg tgggcatggc agttggtata
      C  F  R   P  A  G   H  Q  T   G  I  P   T  V  G   M  A  V  G  I
```

Fig. 2B

```
2461 ctgctgacca cccttctggt gattggtata attttagcag ttgtgtttat ccgcataaag
      L  L  T   T  L  L   V  I  G  I   I  L  A   V  V  F   I  R  I  K
2521 aaggataaag gcaaagataa tgttgaaagt gctcaagcat ctgaagtcaa acctctgaga
      K  D  K   G  K  D   N  V  E  S   A  Q  A   S  E  V   K  P  L  R
2581 agctgaattt gaaaaggaat gtttgaattt atatagcaag tgctatttca gcaacaacca
      S
2641 tctcatccta ttacttttca tctaacgtgc attataattt tttaaacaga tattccctct
2701 tgtcctttaa tatttgctaa atatttcttt tttgaggtgg agtcttgctc tgtcgcccag
2761 gctggagtac agtggtgtga tcccagctca ctgcaacctc cgcctcctgg gttcacatga
2821 ttctcctgcc tcagcttcct aagtagctgg gtttacaggc acccaccacc atgcccagct
2881 aatttttgta tttttaatag agacggggtt tcgccatttg gccaggctgg tcttgaactc
2941 ctgacgtcaa gtgatctgcc tgccttggtc tcccaataca ggcatgaacc actgcaccca
3001 cctacttaga tatttcatgt gctatagaca ttagagagat ttttcatttt tccatgacat
3061 ttttcctctc tgcaaatggc ttagctactt gtgtttttcc cttttggggc aagacagact
3121 cattaaatat tctgtacatt ttttctttat caaggagata tatcagtgtt gtctcataga
3181 actgcctgga ttccatttat gtttttctg attccatcct gtgtcccctt catccttgac
3241 tcctttggta tttcactgaa tttcaaacat ttgtcagaga agaaaaaagt gaggactcag
3301 gaaaaataaa taaataaaag aacagccttt tgcggccgcg aattc
```

Fig. 2C

20	40	60
MARKKFSGLEISLIVLFVIV	TIIAIALIVVLATKTPAVDE	ISDSTSTPATTRVTTNPSDS
80	100	120
GKCPNVLNDPVNVRINCIPE	QFPTEGICAQRGCCWRPWND	SLIPWCFFVDNHGYNVQDMT
140	160	180
TTSIGVEAKLNRIPSPTLFG	NDINSVLFTTQNQTPNRFRF	KITDPNNRRYEVPHQYVKEF
200	220	240
TGPTVSDTLYDVKVAQNPFS	IQVIRKSNGKTLFDTSIGPL	VYSDQYLQISARLPSDYIYG
260	280	300
IGEQVHKRFRHDLSWKTWPI	FTRDQLPGDNNNNLYGHQTF	FMCIEDTSGKSFGVFLMNSN
320	340	360
AMEIFIQPTPIVTYRVTGGI	LDFYILLGDTPEQVVQQYQQ	LVGLPAMPAYWNLGFQLSRW
380	400	420
NYKSLDVVKEVVRRNREAGI	PFDTQVTDIDYMEDKKDFTY	DQVAFNGLPQFVQDLHDHGQ
440	460	480
KYVIILDPAISIGRRANGTT	YATYERGNTQHVWINESDGS	TPIIGEVWPGLTVYPDFTNP
500	520	540
NCIDWWANECSIFHQEVQYD	GLWIDMNEVSSFIQGSTKGC	NVNKLNYPPFTPDILDKLMY
560	580	600
SKTICMDAVQNWGKQYDVHS	LYGYSMAIATEQAVQKVFPN	KRSFILTRSTFAGSGRHAAH
620	640	660
WLGDNTASWEQMEWSITGML	EFSLFGIPLVGADICGFVAE	TTEELCRRWMQLGAFYPFSR
680	700	720
NHNSDGYEHQDPAFFGQNSL	LVKSSRQYLTIRYTLLPFLY	TLFYKAHVFGETVARPVLHE
740	760	780
FYEDTNSWIEDTEFLWGPAL	LITPVLKQGADTVSAYIPDA	IWYDYESGAKRPWRKQRVDM
800	820	840
YLPADKIGLHLRGGYIIPIQ	EPDVTTTASRKNPLGLIVAL	GENNTAKGDFFWDDGETKDT
860	880	900
IQNGNYILYTFSVSNNTLDI	VCTHSSYQEGTTLAFQTVKI	LGLTDSVTEVRVAENNQPMN
920	940	960
AHSNFTYDASNQVLLIADLK	LNLGRNFSVQWNQIFSENER	FNCYPDADLATEQKCTQRGC
980	1000	1020
VWRTGSSLSKAPECYFPRQD	NSYSVNSARYSSMGITADLQ	LNTANARIKLPSDPISTLRV
1040	1060	1080
EVKYHKNDMLQFKIYDPQKK	RYEVPVPLNIPTTPISTYED	RLYDVEIKENPFGIQIRRRS
1100	1120	1140
SGRVIWDSWLPGFAFNDQFI	QISTRLPSEYIYGFGEVEHT	AFKRDLNWNTWGMFTRDQPP
1160	1180	1200
GYKLNSYGFHPYYMALEEEG	NAHGVFLLNSNAMDVTFQPT	PALTYRTVGGILDFYMFLGP
1220	1240	1260
TPQVATKQYHEVIGHPVMPA	YWALGFQLCRYGYANTSEVR	ELYDAMVAANIPYDVQYTDI

Fig. 3A

1280	1300	1320
DYMERQLDFTIGEAFQDLPQ	FVDKIRGEGMRYIIILDPAI	SGNETKTYPAFERGQQNDVF
1340	1360	1380
VKWPNTNDICWAKVWPDLPN	ITIDKTLTEDEAVNASRAHV	AFPDFFRTSTAEWWAREIVD
1400	1420	1440
FYNEKMKFDGLWIDMNEPSS	FVNGTTTNQCRNDELNYPPY	FPELTKRTDGLHFRTICMEA
1460	1480	1500
EQILSDGTSVLHYDVHNLYG	WSQMKPTHDALQKTTGKRGI	VISRSTYPTSGRWGGHWLGD
1520	1540	1560
NYARWDNMDKSIIGMMEFSL	FGISYTGADICGFFNNSEYH	LCTRWMQLGAFYPYSRNHNI
1580	1600	1620
ANTRRQDPASWNETFAEMSR	NILNIRYTLLPYFYTQMHEI	HANGGTVIRPLLHEFFDEKP
1640	1660	1680
TWDIFKQFLWGPAFMVTPVL	EPYVQTVNAYVPNARWFDYH	TGKDIGVRGQFQTFNASYDT
1700	1720	1740
INLHVRGGHILPCQEPAQNT	FYSRQKHMKLIVAADDNQMA	QGSLFWDDGESIDTYERDLY
1760	1780	1800
LSVQFNLNQTTLTSTILKRG	YINKSETRLGSLHVWGKGTT	PVNAVTLTYNGNKNSLPFNE
1820	1827	
DTTNMILRIDLTTHNVTLEE	PIEINWS	

Fig. 3B

```
  1 gccttactgc aggaaggcac tccgaagaca taagtcggtg agacatggct gaagataaaa
                                                   M  A   E  D  K
 61 gcaagagaga ctccatcgag atgagtatga agggatgcca gacaaacaac gggtttgtcc
     S  K  R   D  S  I  E  M  S  M   K  G  C   Q  T  N  N   G  F  V
121 ataatgaaga cattctggag cagaccccgg atccaggcag ctcaacagac aacctgaagc
     H  N  E   D  I  L  E  Q  T  P   D  P  G   S  S  T  D   N  L  K
181 acagcaccag gggcatcctt ggctcccagg agcccgactt caagggcgtc cagccctatg
     H  S  T   R  G  I  L  G  S  Q   E  P  D   F  K  G  V   Q  P  Y
241 cggggatgcc caaggaggtg ctgttccagt tctctggcca ggcccgctac cgcatacctc
     A  G  M   P  K  E  V  L  F  Q   F  S  G   Q  A  R  Y   R  I  P
301 gggagatcct cttctggctc acagtggctt ctgtgctggt gctcatcgcg gccaccatag
     R  E  I   L  F  W  L  T  V  A   S  V  L   V  L  I  A   A  T  I
361 ccatcattgc cctctctcca aagtgcctag actggtggca ggagggcccc atgtaccaga
     A  I  I   A  L  S  P  K  C  L   D  W  W   Q  E  G  P   M  Y  Q
421 tctacccaag gtctttcaag gacagtaaca aggatgggaa cggagatctg aaaggtattc
     I  Y  P   R  S  F  K  D  S  N   K  D  G   N  G  D  L   K  G  I
481 aagataaact ggactacatc acagctttaa atataaaaac tgtttggatt acttcatttt
     Q  D  K   L  D  Y  I  T  A  L   N  I  K   T  V  W  I   T  S  F
541 ataaatcgtc ccttaaagat ttcagatatg gtgttgaaga tttccgggaa gttgatccca
     Y  K  S   S  L  K  D  F  R  Y   G  V  E   D  F  R  E   V  D  P
601 tttttggaac gatggaagat tttgagaatc tggttgcagc catacatgat aaaggtttaa
     I  F  G   T  M  E  D  F  E  N   L  V  A   A  I  H  D   K  G  L
661 aattaatcat cgatttcata ccaaaccaca cgagtgataa acatatttgg tttcaattga
     K  L  I   I  D  F  I  P  N  H   T  S  D   K  H  I  W   F  Q  L
721 gtcggacacg gacaggaaaa tatactgatt attatatctg gcatgactgt acccatgaaa
     S  R  T   R  T  G  K  Y  T  D   Y  Y  I   W  H  D  C   T  H  E
781 atggcaaaac cattccaccc aacaactggt taagtgtgta tggaaactcc agttggcact
     N  G  K   T  I  P  P  N  N  W   L  S  V   Y  G  N  S   S  W  H
841 ttgacgaagt gcgaaaccaa tgttattttc atcagtttat gaaagagcaa cctgatttaa
     F  D  E   V  R  N  Q  C  Y  F   H  Q  F   M  K  E  Q   P  D  L
901 atttccgcaa tcctgatgtt caagaagaaa taaaagaaat tttacggttc tggctcacaa
     N  F  R   N  P  D  V  Q  E  E   I  K  E   I  L  R  F   W  L  T
961 agggtgttga tggttttagt ttggatgctg ttaaattcct cctagaagca aagcacctga
     K  G  V   D  G  F  S  L  D  A   V  K  F   L  L  E  A   K  H  L
```

Fig. 4A

```
1021 gagatgagat ccaagtaaat aagacccaaa tcccggacac ggtcacacaa tactcggagc
      R  D  E   I  Q  V  N   K  T  Q   I  P  D   T  V  T  Q   Y  S  E
1081 tgtaccatga cttcaccacc acgcaggtgg gaatgcacga cattgtccgc agcttccggc
      L  Y  H   D  F  T  T   T  Q  V   G  M  H   D  I  V  R   S  F  R
1141 agaccatgga ccaatacagc acggagcccg gcagatacag gttcatgggg actgaagcct
      Q  T  M   D  Q  Y  S   T  E  P   G  R  Y   R  F  M  G   T  E  A
1201 atgcagagag tattgacagg accgtgatgt actatggatt gccatttatc caagaagctg
      Y  A  E   S  I  D  R   T  V  M   Y  Y  G   L  P  F  I   Q  E  A
1261 attttccctt caacaattac ctcagcatgc tagacactgt ttctgggaac agcgtgtatg
      D  F  P   F  N  N  Y   L  S  M   L  D  T   V  S  G  N   S  V  Y
1321 aggttatcac atcctggatg gaaaacatgc cagaaggaaa atggcctaac tggatgattg
      E  V  I   T  S  W  M   E  N  M   P  E  G   K  W  P  N   W  M  I
1381 gtggaccaga cagttcacgg ctgacttcgc gtttggggaa tcagtatgtc aacgtgatga
      G  G  P   D  S  S  R   L  T  S   R  L  G   N  Q  Y  V   N  V  M
1441 acatgcttct tttcacactc cctggaactc ctataactta ctatggagaa gaaattggaa
      N  M  L   L  F  T  L   P  G  T   P  I  T   Y  Y  G  E   E  I  G
1501 tgggaaatat tgtagccgca aatctcaatg aaagctatga tattaatacc cttcgctcaa
      M  G  N   I  V  A  A   N  L  N   E  S  Y   D  I  N  T   L  R  S
1561 agtcaccaat gcagtgggac aatagttcaa atgctggttt ttctgaagct agtaacacct
      K  S  P   M  Q  W  D   N  S  S   N  A  G   F  S  E  A   S  N  T
1621 ggttacctac caattcagat taccacactg tgaatgttga tgtccaaaag actcagccca
      W  L  P   T  N  S  D   Y  H  T   V  N  V   D  V  Q  K   T  Q  P
1681 gatcggcttt gaagttatat caagatttaa gtctacttca tgccaatgag ctactcctca
      R  S  A   L  K  L  Y   Q  D  L   S  L  L   H  A  N  E   L  L  L
1741 acaggggctg gttttgccat ttgaggaatg acagccacta tgttgtgtac acaagagagc
      N  R  G   W  F  C  H   L  R  N   D  S  H   Y  V  V  Y   T  R  E
1801 tggatggcat cgacagaatc tttatcgtgg ttctgaattt tggagaatca acactgttaa
      L  D  G   I  D  R  I   F  I  V   V  L  N   F  G  E  S   T  L  L
1861 atctacataa tatgatttcg ggccttcccg ctaaaataag aataaggtta agtaccaatt
      N  L  H   N  M  I  S   G  L  P   A  K  I   R  I  R  L   S  T  N
1921 ctgccgacaa aggcagtaaa gttgatacaa gtggcatttt tctggacaag ggagagggac
      S  A  D   K  G  S  K   V  D  T   S  G  I   F  L  D  K   G  E  G
1981 tcatctttga acacaacacg aagaatctcc ttcatcgcca aacagctttc agagatagat
      L  I  F   E  H  N  T   K  N  L   L  H  R   Q  T  A  F   R  D  R
2041 gctttgtttc caatcgagca tgctattcca gtgtactgaa catactgtat acctcgtgtt
      C  F  V   S  N  R  A   C  Y  S   S  V  L   N  I  L  Y   T  S  C
2101 aggcaccttt atgaagagat gaagacactg gcatttcagt gggattgtaa gcatttgtaa
2161 tagcttcatg tacagcatgc tgcttggtga acaatcatta attcttcgat atttctgtag
2221 cttgaatgta accgctttaa gaaaggttct caaatgtttt gaaaaaaata aaatgtttaa
2281 aagt
```

Fig. 4B

Expression of Phage Inserts as GST Fusion

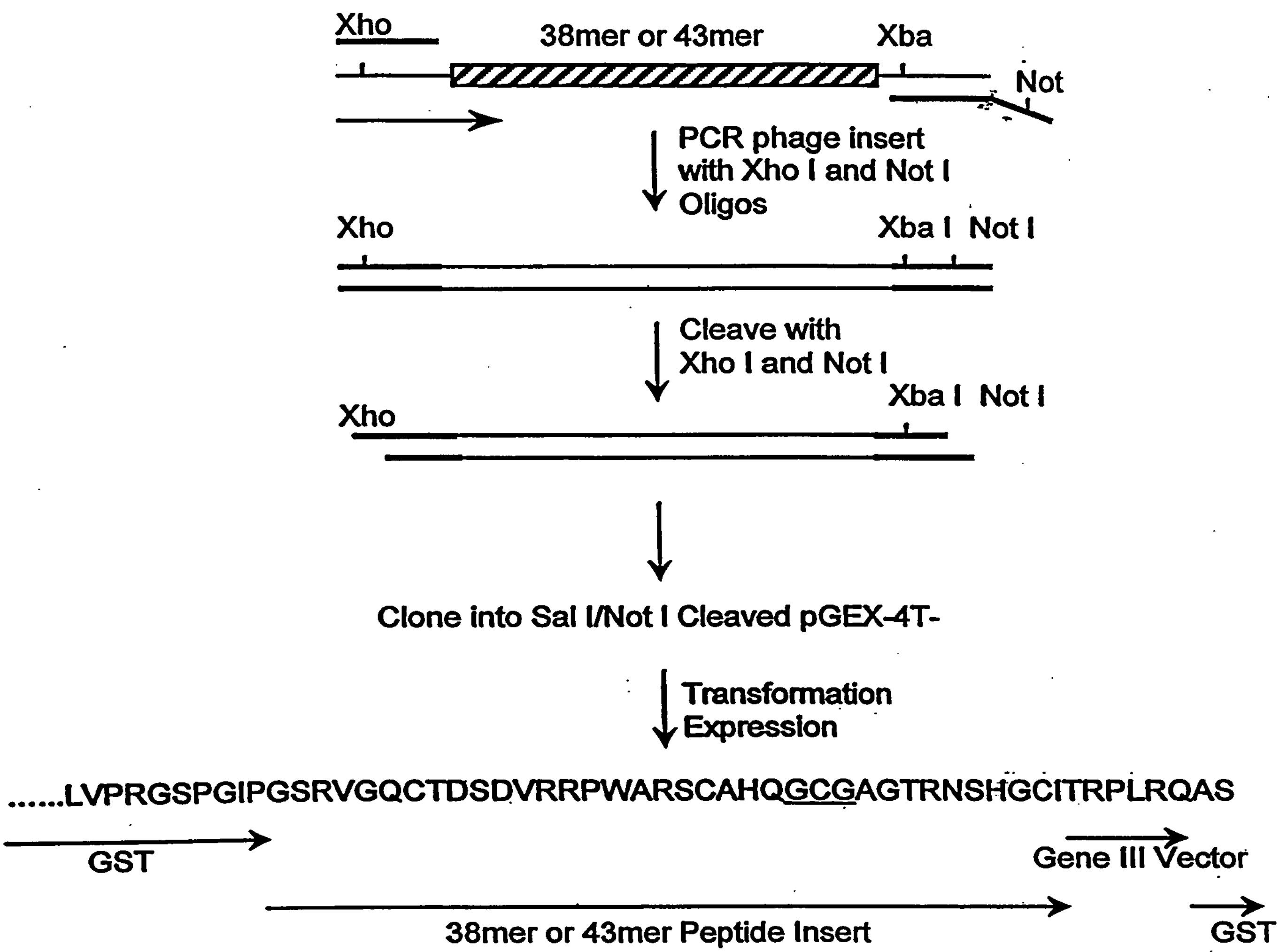


Fig. 5A

P31	Clone #
SARDSGPAEDGSRAVRLNGVENANTRKSSRSNPRGRRHP	
SARDSGPAEDGSRAVRLNG	101
DGSRAVRLNGVENANTRKSSR	102
ENANTRKSSRSNPRGRRHP	103
TRKSSRSNPRG	119

Pax2	Clone #
STPPSREAYSRPYSVDSDSDTNAKHSSHNRRLRTRSRPN	
STPPSREAYSRPYSVDSDSD	104
SRPYSVDSDSDTNAKHSSHNR	105
TNAKHSSHNRRLRTRSRPN	106

DCX8	Clone #
RYKHDIGCDAGVDKKSSSVRGGCGAHSSPPRAGRGPRGTMVSRL	
RYKHDIGCDAGVDKKSSSVRGGCG	107
GCDAGVDKKSSSVRGGCGAHSSPPRA	108
GAHSSPPRAGRGPRGTMVSRL	109

Fig. 5B

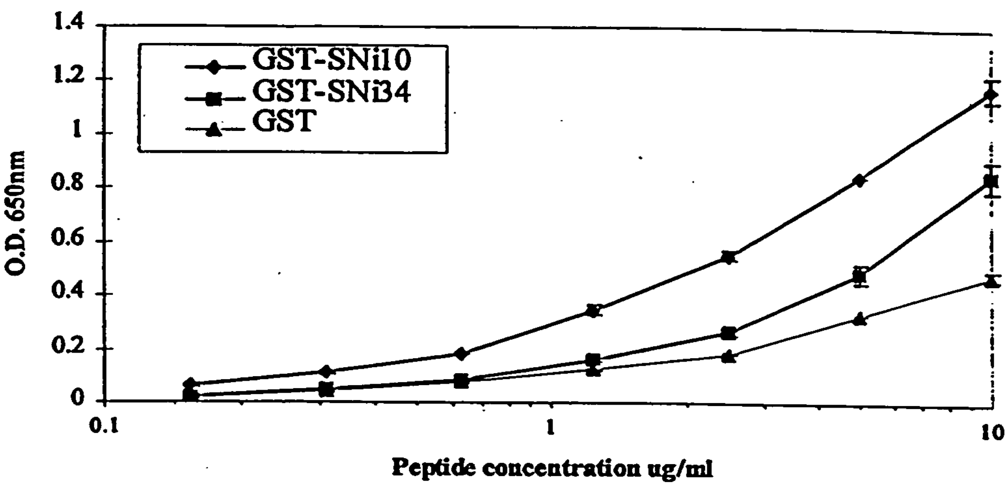
P31 1 10 20 30	Clone #
SARDSGPAEDGSRAVRLNGVEN<u>A</u>NTRKSSRSNPRGRRHP	
EN<u>A</u>NTRKSSRSNPRGRRHP	103
EN<u>A</u>NTRKSSR	110
TRKSSRSNPRG	119
RKSSRSNPRG	111
SNPRGRRHP	112

Pax2 1 10 20 30	Clone #
STPPSREAYSRPYSVDSDSDTN<u>A</u>KHSSHNRRLRTRSRPN	
TN<u>A</u>KHSSHNRRLRTRSRPN	106
TN<u>A</u>KHSSHN	113
SSHNRRLRTR	114
RRLRTRSRPN	115

SNi10 1 10 20 30	Cl ne #
RVGQCTDSDVRRPWARSCAHQ<u>GCG</u>AGTRNSHGCITRPLRQASAH	
RVGQCTDSDVRRPWARSCA	116
VRRPWARSCAHQ<u>GCG</u>AGTRNS	117
GTRNSHGCITRPLRQASAH	118

Fig. 5C

A


GST-SNi10
GST-SNi34
GST
O.D. 650nm
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.1
1
10
Peptide concentration ug/ml

B

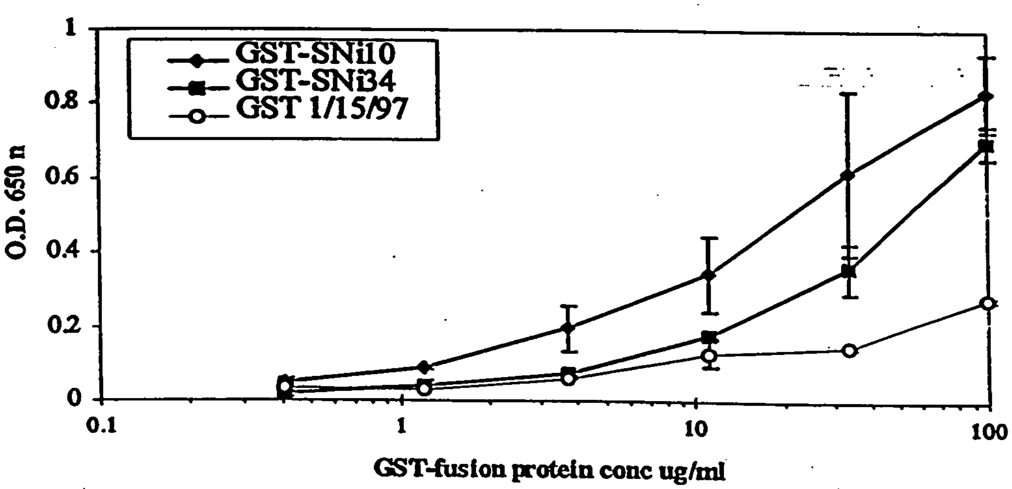

GST-SNi10
GST-SNi34
GST 1/15/97
O.D. 650 n
1
0.8
0.6
0.4
0.2
0
0.1
1
10
100
GST-fusion protein conc ug/ml

Fig. 6

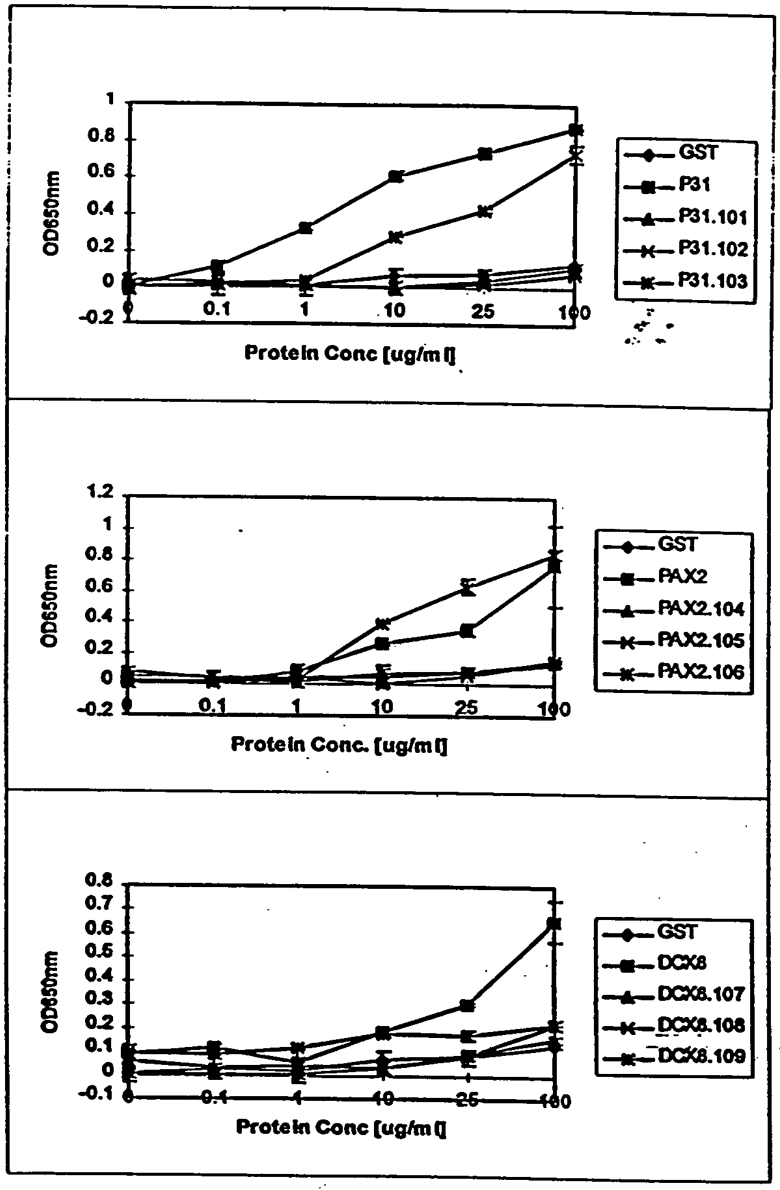
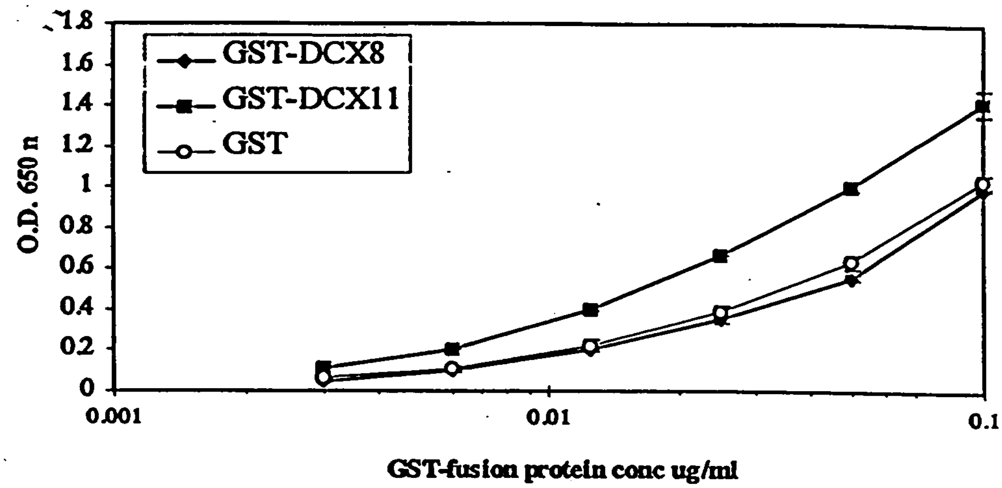

OD650nm
1
0.8
0.6
0.4
0.2
0
-0.2
0
0.1
1
10
25
100
Protein Conc [ug/ml]
GST
P31
P31.101
P31.102
P31.103
OD650nm
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
0.1
1
10
25
100
Protein Conc. [ug/ml]
GST
PAX2
PAX2.104
PAX2.105
PAX2.106
OD650nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
0
0.1
1
10
25
100
Protein Conc [ug/ml]
GST
DCX8
DCX8.107
DCX8.108
DCX8.109

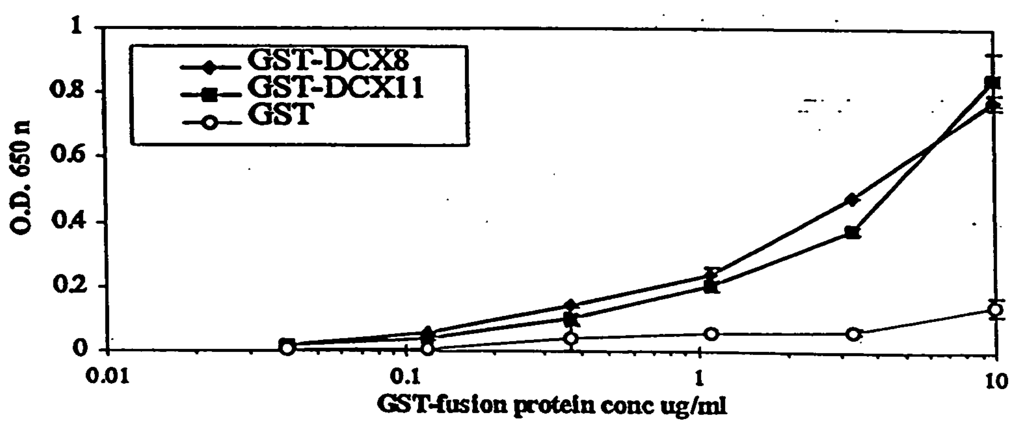
Fig. 7A-C

D



1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
O.D. 650 n
GST-DCX8
GST-DCX11
GST
0.001
0.01
0.1
GST-fusion protein conc ug/ml

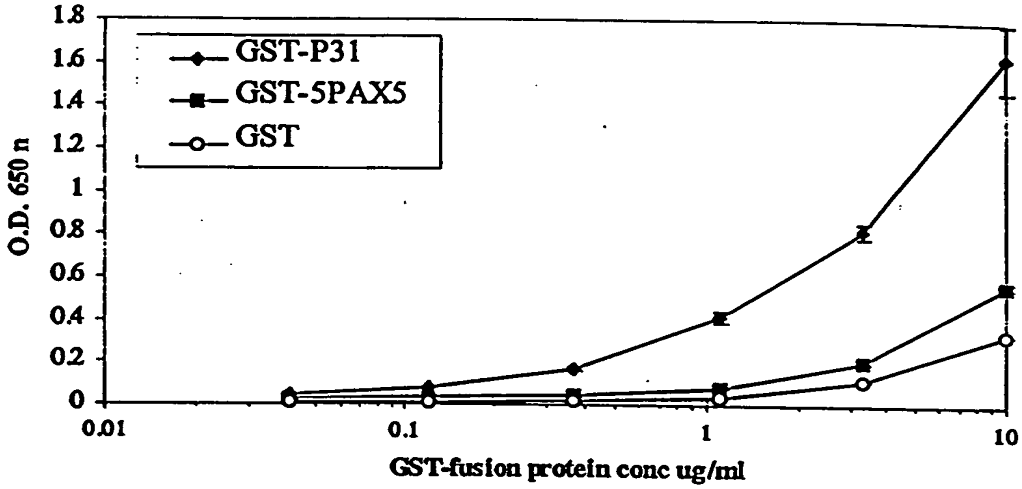
E



1
0.8
0.6
0.4
0.2
0
O.D. 650 n
GST-DCX8
GST-DCX11
GST
0.01
0.1
1
10
GST-fusion protein conc ug/ml

Fig. 7 D-E

F


1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
O.D. 650 n
GST-P31
GST-5PAX5
GST
0.01
0.1
1
10
GST-fusion protein conc ug/ml

G

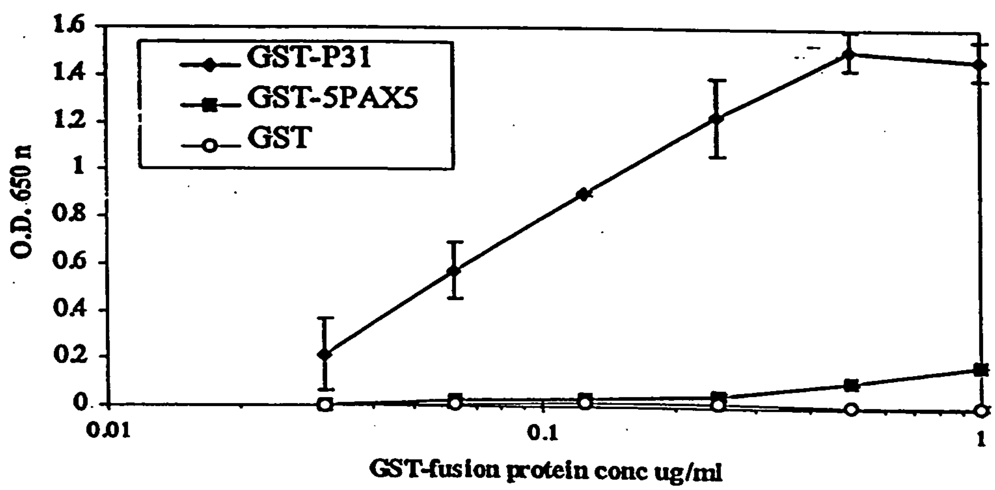
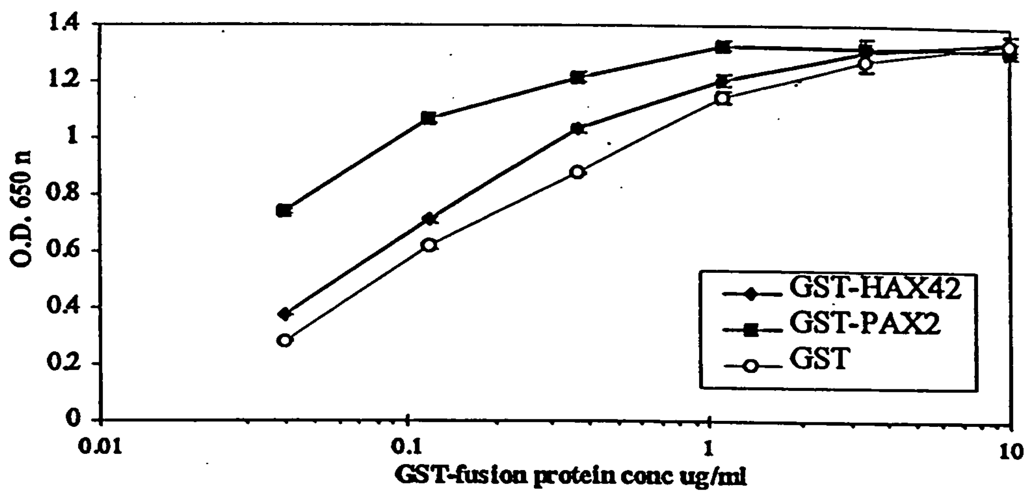

1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0.0
O.D. 650 n
GST-P31
GST-5PAX5
GST
0.01
0.1
1
GST-fusion protein conc ug/ml

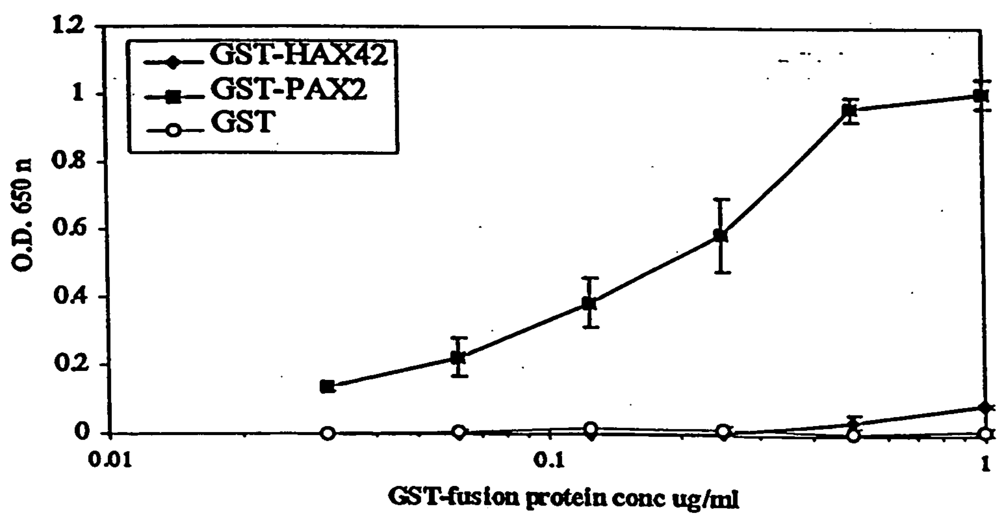
Fig. 7 F-G

H



1.4
1.2
1
0.8
0.6
0.4
0.2
0
O.D. 650 n
0.01
0.1
1
10
GST-fusion protein conc ug/ml
GST-HAX42
GST-PAX2
GST

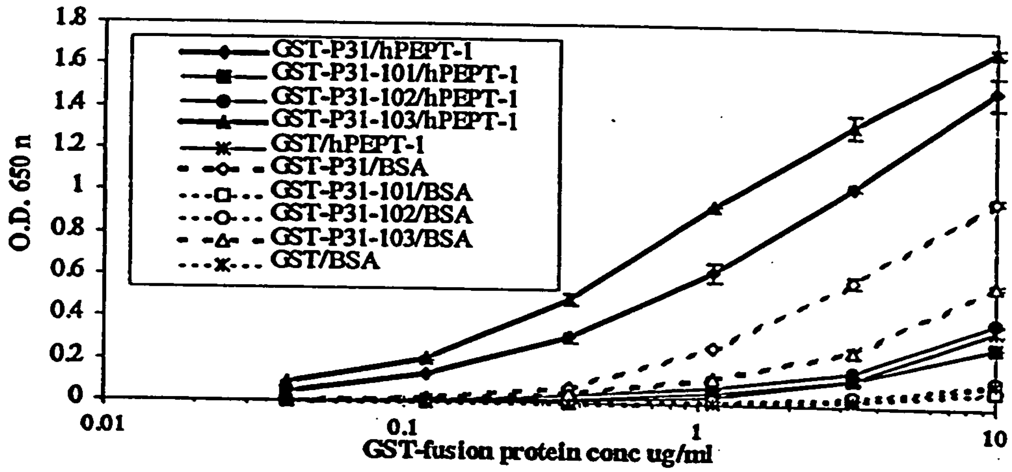
I



1.2
1
0.8
0.6
0.4
0.2
0
O.D. 650 n
0.01
0.1
1
GST-fusion protein conc ug/ml
GST-HAX42
GST-PAX2
GST

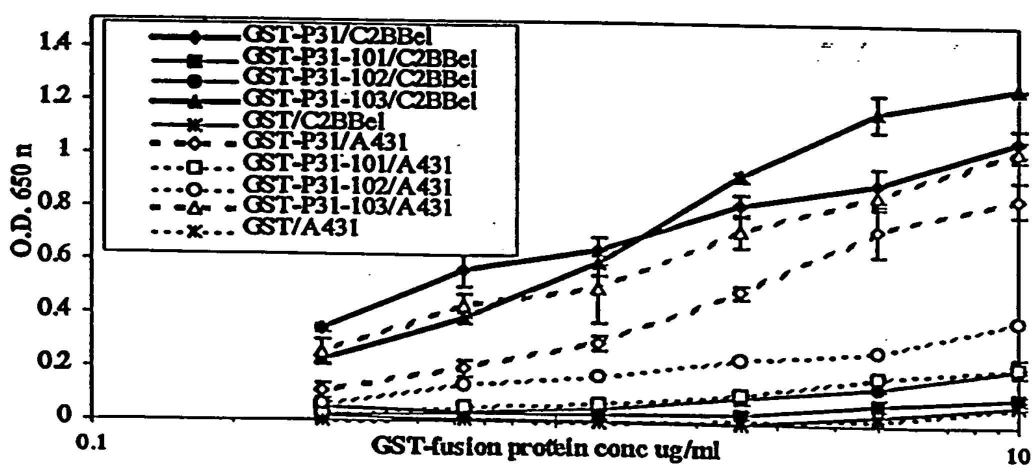
Fig. 7 H-I

J



1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
O.D. 650 n
0.01
0.1
1
10
GST-fusion protein conc ug/ml
GST-P31/hPEPT-1
GST-P31-101/hPEPT-1
GST-P31-102/hPEPT-1
GST-P31-103/hPEPT-1
GST/hPEPT-1
GST-P31/BSA
GST-P31-101/BSA
GST-P31-102/BSA
GST-P31-103/BSA
GST/BSA

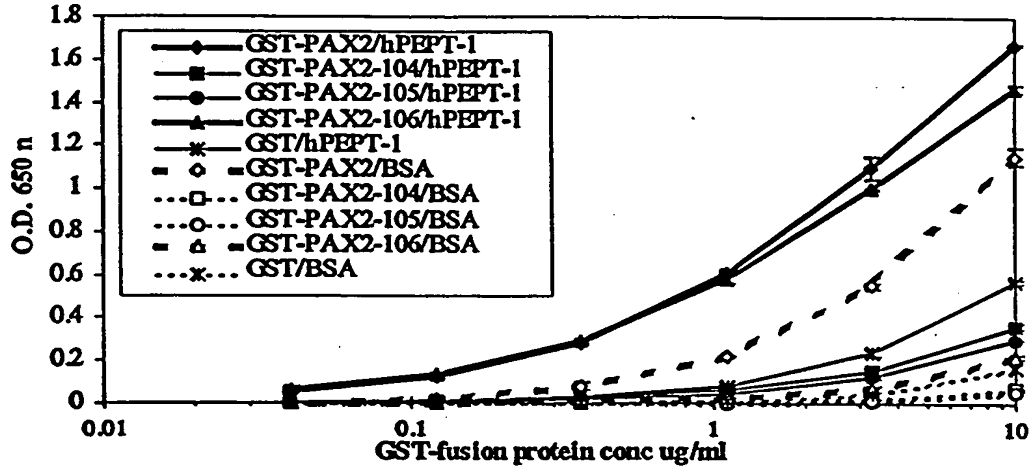
K



1.4
1.2
1
0.8
0.6
0.4
0.2
0
O.D. 650 n
0.1
10
GST-fusion protein conc ug/ml
GST-P31/C2BBel
GST-P31-101/C2BBel
GST-P31-102/C2BBel
GST-P31-103/C2BBel
GST/C2BBel
GST-P31/A431
GST-P31-101/A431
GST-P31-102/A431
GST-P31-103/A431
GST/A431

Fig. 7 J-K

L


GST-PAX2/hPEPT-1
GST-PAX2-104/hPEPT-1
GST-PAX2-105/hPEPT-1
GST-PAX2-106/hPEPT-1
GST/hPEPT-1
GST-PAX2/BSA
GST-PAX2-104/BSA
GST-PAX2-105/BSA
GST-PAX2-106/BSA
GST/BSA
O.D. 650 n
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.01
0.1
1
10
GST-fusion protein conc ug/ml

M

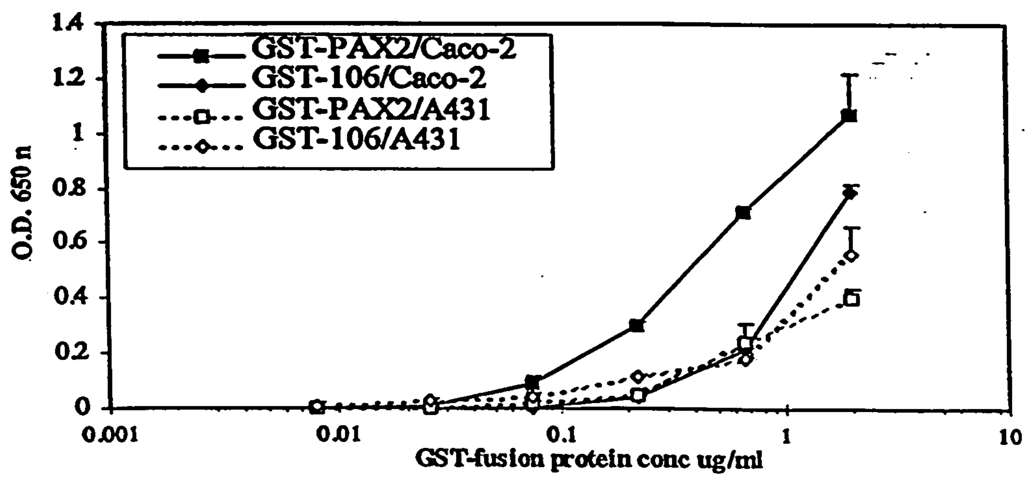
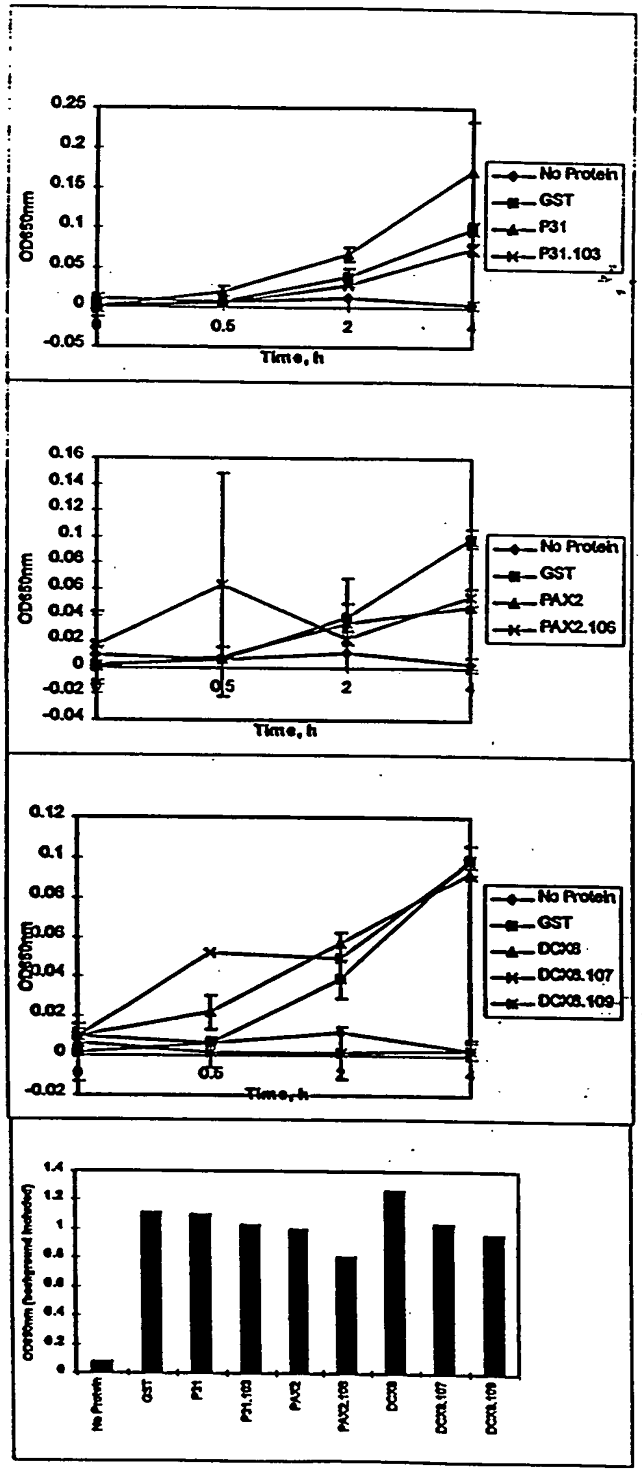
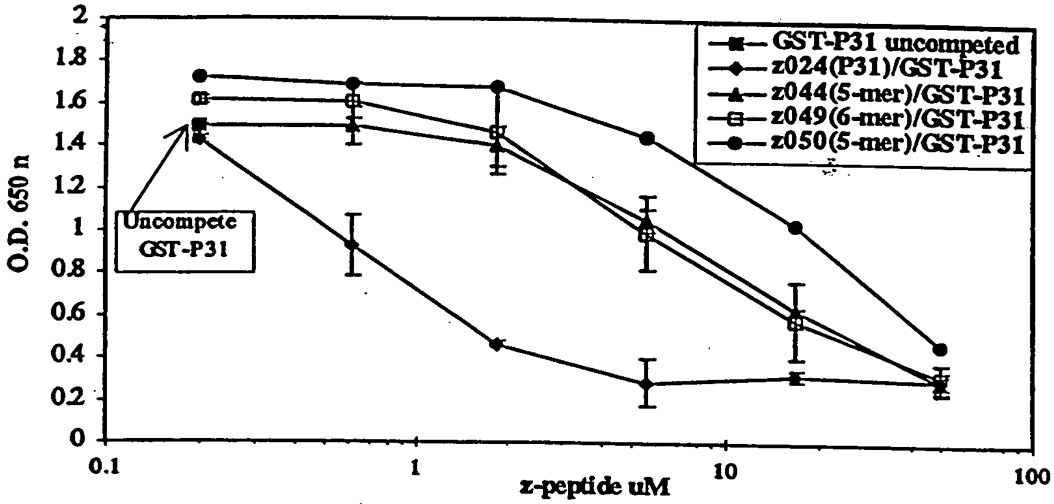

GST-PAX2/Caco-2
GST-106/Caco-2
GST-PAX2/A431
GST-106/A431
O.D. 650 n
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.001
0.01
0.1
1
10
GST-fusion protein conc ug/ml

Fig. 7 L-M

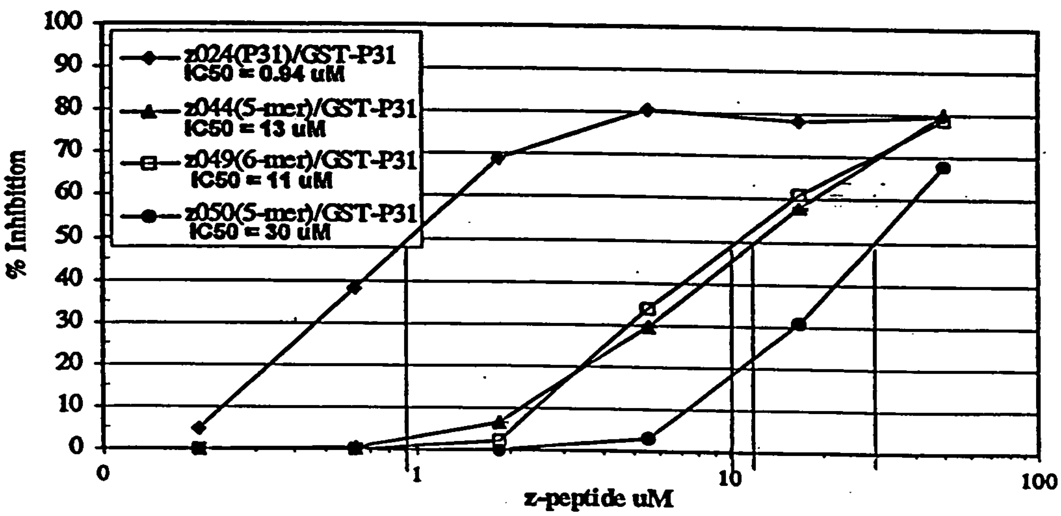

OD650nm
Time, h
No Protein
GST
P31
P31.103
PAX2
PAX2.106
DCX8
DCX8.107
DCX8.109
OD650nm (background included)

Figs. 8 A-D



GST-P31 uncompeted
z024(P31)/GST-P31
z044(5-mer)/GST-P31
z049(6-mer)/GST-P31
z050(5-mer)/GST-P31
Uncompete GST-P31
O.D. 650 n
z-peptide uM
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.1
1
10
100

A



z024(P31)/GST-P31 IC50 = 0.94 uM
z044(5-mer)/GST-P31 IC50 = 13 uM
z049(6-mer)/GST-P31 IC50 = 11 uM
z050(5-mer)/GST-P31 IC50 = 30 uM
% Inhibition
z-peptide uM
100
90
80
70
60
50
40
30
20
10
0
0
1
10
100

B

Fig. 9

865T50"8Z96Z060

P31

Peptide Name	Sequence (1 · 10 · 20 · 30 · 40)	pI	IC_{50}	GST/C2BBe1	
ELAN024 (P31)	SARDSGPAEDGSRAVRLNGVENANTRKSSRSNPRGRRHPG	11.88	0.5-2.2	+++	
101	SARDSGPAEDGSRAVRLNG				-
102	DGSRAVRLNGVENANTRKSSR			-	
103	ENANTRKSSRSNPRGRRHP				++
110	ENANTRKSSR			-	
111	RKSSRSNPRG			-	
112	SNPRGRRHP				-
119	TRKSSRSNPRG				
Z28	ZENANTRKSSRSNPRGRRHPG	12.28	0.5-1.7		
Z29	ZTRKSSRSNPRG	12.40	5.5-15		
Z30	ZENANTRKSSRSNPRG	11.81	>50		
Z31	ZTRKSSRSNPRGRRHPG	12.70	0.6-3.2		
Z39	ZENANTRKSSR	10.89	>50		
Z40	ZSNPRGRRHPG	12.40	5.9-29		
Z41	ZENANT	3.75	>50		
Z42	ZANTRKS	11.05	>50		
Z43	ZTRKSS	11.05	>50		
Z44	ZRKSSR	12.11	13->50		
Z45	ZKSSRSN	11.05	40-48		
Z46	ZSSRSNPG	10.04	>50		
Z47	ZRSNPRG	12.40	>50		
Z48	ZSNPRG	10.04	>50		
Z49	ZPRGRRH	12.40	11-20		
Z50	ZRRHPG	12.10	30		
Z51 (HepC core)	ZKSSRGN	12.40	>50		
Z52 (HepC P26664)	ZKTSERSQPRGRRQPG	12.10	9.8		
Z53	ZTrKSSrSNPrGrrHPG		1.6		
Z54	ZTRKSSrSNPRGrRHPG		1.6		
Z21 (HAX42)	SDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIPT	11.27	1.7		

Fig. 10A

PAX2

Peptide Name	Sequence	pI	IC$_{50}$	GST/C2BBe1
	1 10 20 30 40			
ELAN018 (PAX2)	STPPSREAYSRPYSVDSDSDTNAKHSSHNRRLRTRSRPNG	10.88	0.6-0.9, 1	+++
104	STPPSREAYSRPYSVDSDSD			-
105	SRPYSVDSDSDTNAKHSSHNR			-
106	TNAKHSSHNRRLRTRSRPN			++
113	TNAKHSSHN			-
114	SSHNRRLRTR			+/-
115	RRLRTRSRPN			+/-
Z32	ZTNAKHSSHNRRLRTRSRPN	12.7	1.2	
Z33	ZTNAKHSSHNRRLRTR	12.58	1.6	
Z34	ZSSHNRRLRTRSRPN	12.7	1.6, 1.3, 0.68, 1.5	
Z35	ZSSHNRRLRTR	12.58	0.38 - 1.8, 2.7	
Z26	Z SEANLDGRKSRYSSPRRNSSTRPRTSPNSVHARYPSTDHD	10.88	7-8, 3	
Z38	Z SRANTDGRKSRYSSPRRNSSTEPRLSPNSVHARYPSTDHD	10.88	1.7, 0.9	
Z55	ZTNAKHSSHN		42	
Z56	ZRRLRTRSRPN		1.7	
Z57	ZRRLRTRSR		1.9	
Z58	ZRRLRTR		3.4	
Z59	ZrrLrTrSrPN		NOT DONE	
Z73	ZASHNRRLRTR		1.5, 5.5	
Z74	ZSAHNRRLRTR		6.2	
Z75	ZSSANRRLRTR		1.6	
Z76	ZSSHARRLRTR		1.8	
Z77	ZSSHNARLRTR		3.9, 5.2	
Z78	ZSSHNRALRTR		4.5, 4.6	
Z79	ZSSHNRRARTR		1.4	
Z80	ZSSHNRRLATR		3.4, 5.2	
Z81	ZSSHNRRLRAR		2.2	
Z82	ZSSHNRRLRTA		3.4	
Z21 (HAX42)	ZSDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIPT	11.27	0.7	

Fig. 10B

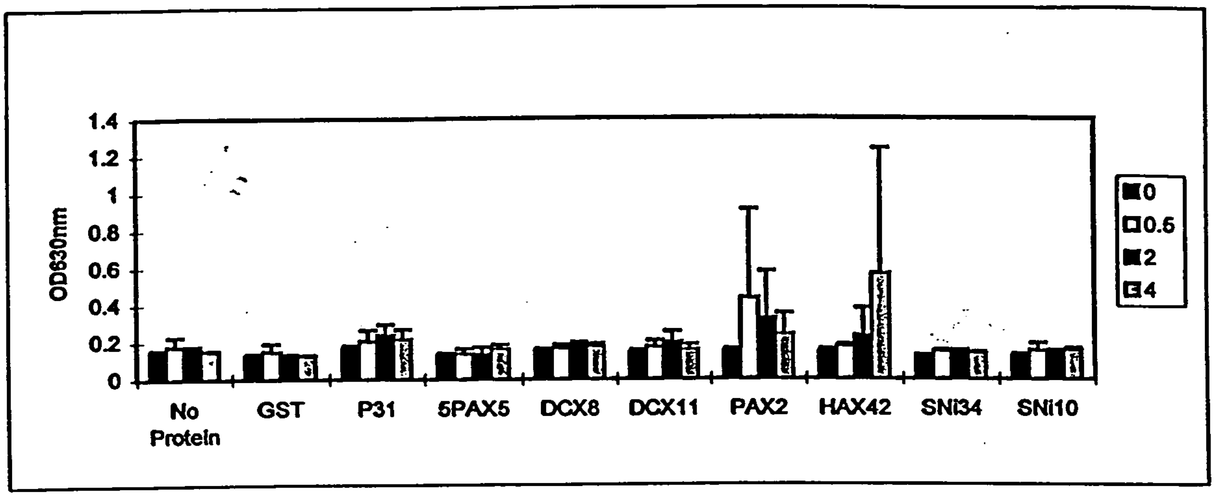
SNi10

Peptide Name	Sequence	pI	IC$_{50}$	GST/C2BBe1
ELAN016(SNi10)	RVGQCTDSDVRRPWARSCAHQGCGAGTRNSHGCITRPLRQASAH	10.19	0.22	++
116	RVGQCTDSDVRRPWARSCA			-
117	VRRPWARSCAHQGCGAGTRNS			+
118	GTRNSHGCITRPLRQASAH			+/-
Z17	ZRVGQCTDSDVRRPWARSCAH	8.66	3.6	
Z16C23	ZCGAGTRNSHGCITRPLRQASAH	9.03	0.7	
Z36	ZVRRPWARSCAHQGCGAGTRNS	11.62	0.27	
Z37	ZCTDSDVRRPWARSC	8.01	3	

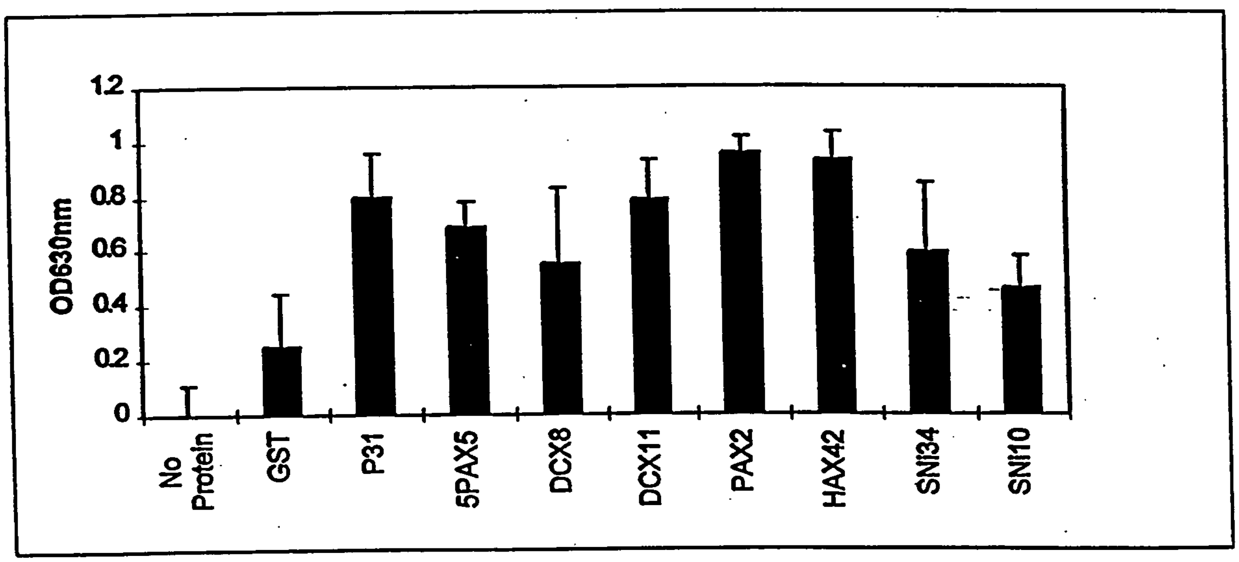
HAX42

Peptide Name	Sequence	pI	IC$_{50}$	GST/C2BBe1
ELAN021(HAX42)	SDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIPT	11.27	5.5	++
ELAN018(PAX2)	STPPSREAYSRPYSVDSDSDTNAKHSSHNRRLRTRSRPNG	10.88	0.23	+++
Z26	ZSEANLDGRKSRYSSPRRNSSTRPRTSPNSVHARYPSTDHD	10.88	<0.2	
Z38	ZSRANTDGRKSRYSSPRRNSSTEPRLSPNSVHARYPSTDHD	10.88	<0.2	
Z34 (PAX2 14mer)	ZSSHNRRLRTRSRPN	12.7	0.33	

Fig. 10C


OD630nm
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
2
4
No Protein
GST
P31
5PAX5
DCX8
DCX11
PAX2
HAX42
SNi34
SNi10

A


OD630nm
1.2
1
0.8
0.6
0.4
0.2
0
No Protein
GST
P31
5PAX5
DCX8
DCX11
PAX2
HAX42
SNi34
SNi10

B

Fig. 11

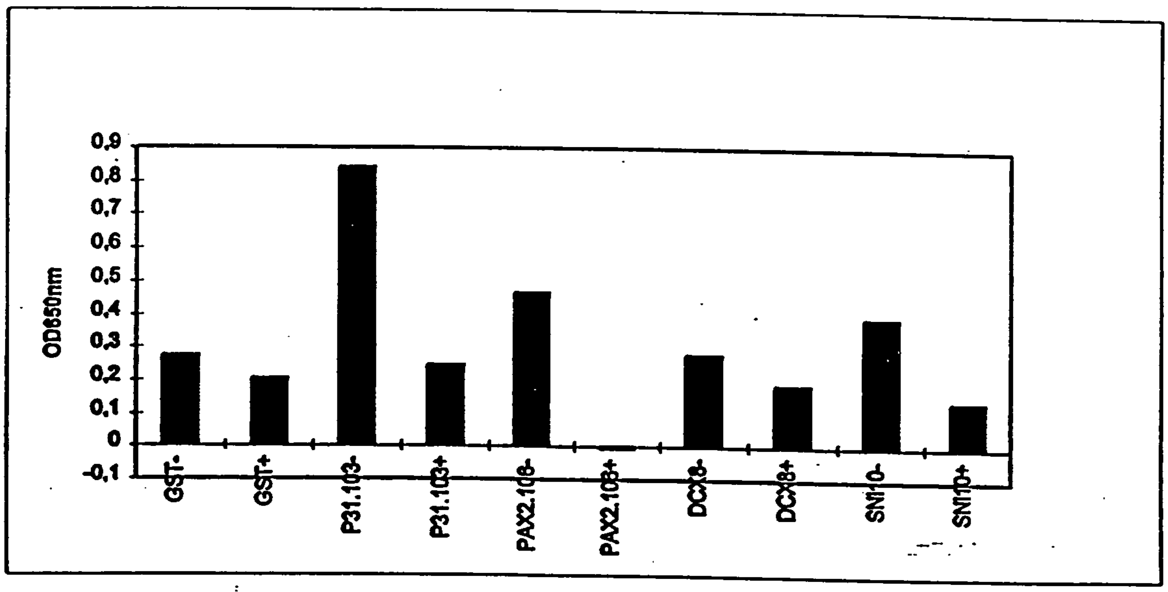
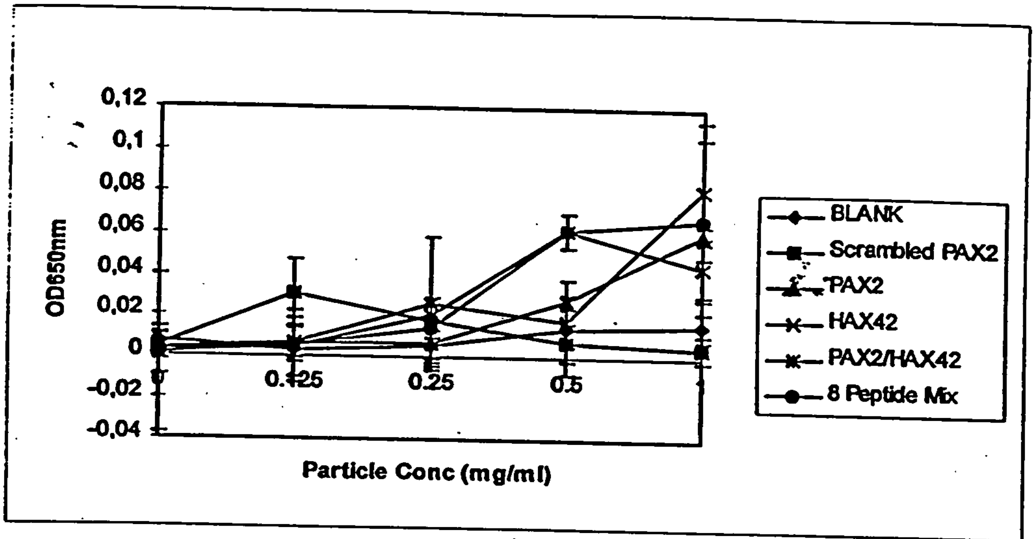
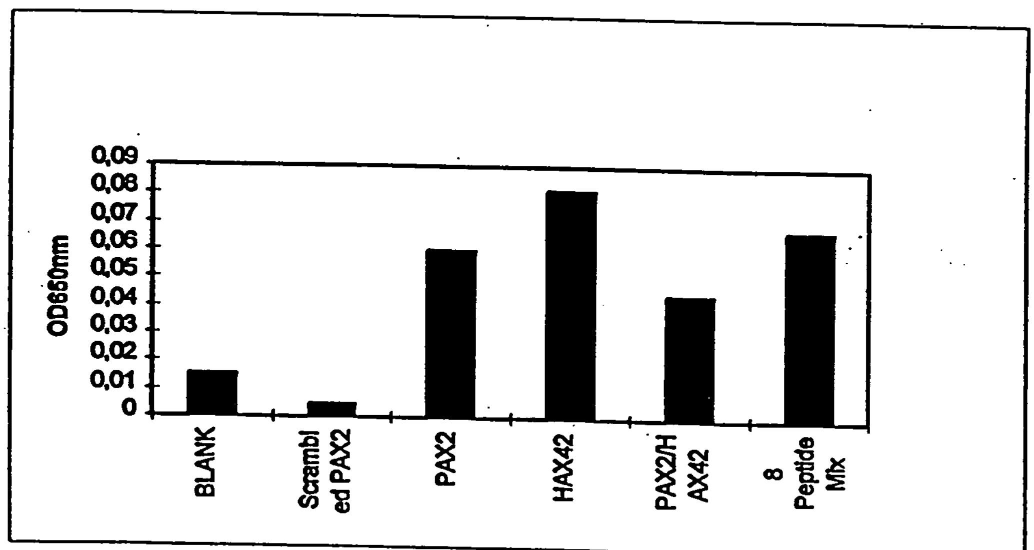

OD650nm
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
GST-
GST+
P31.103-
P31.103+
PAX2.108-
PAX2.108+
DCX8-
DCX8+
SN10-
SN10+

Fig. 12


0,12
0,1
0,08
0,06
0,04
0,02
0
-0,02
-0,04
OD650nm
0
0.125
0.25
0.5
Particle Conc (mg/ml)
BLANK
Scrambled PAX2
PAX2
HAX42
PAX2/HAX42
8 Peptide Mix

A


0,09
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
OD650nm
BLANK
Scrambled PAX2
PAX2
HAX42
PAX2/HAX42
8 Peptide Mix

B

Fig. 13

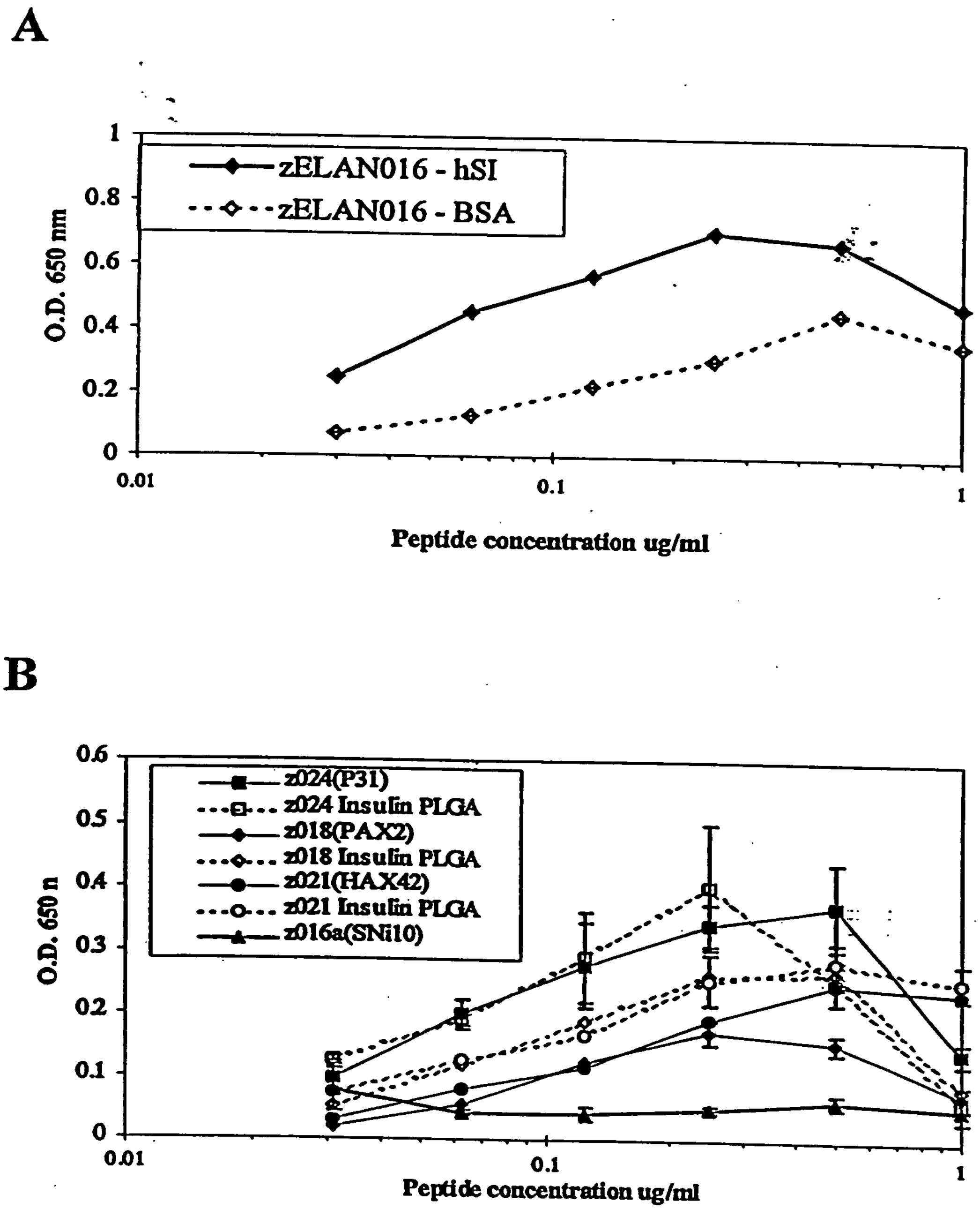
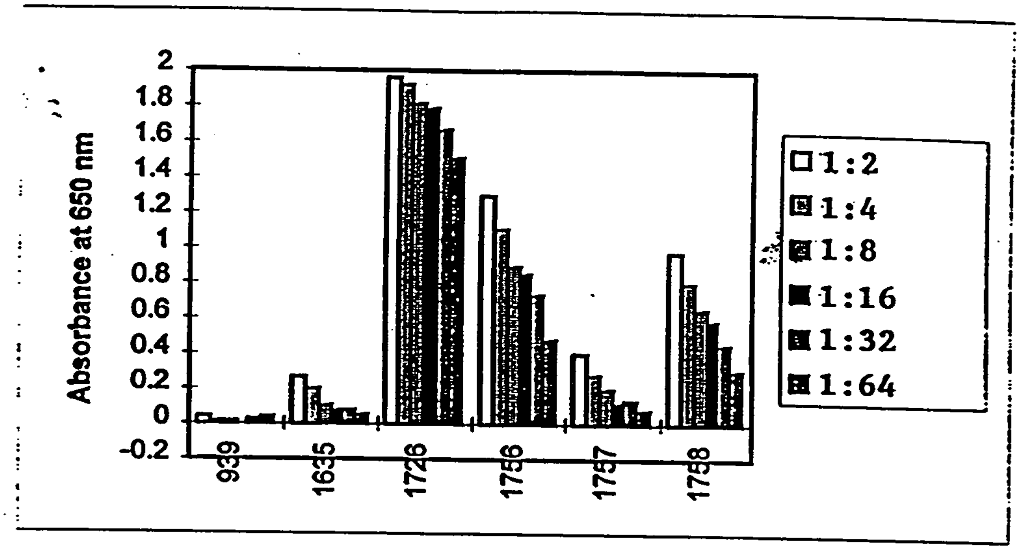
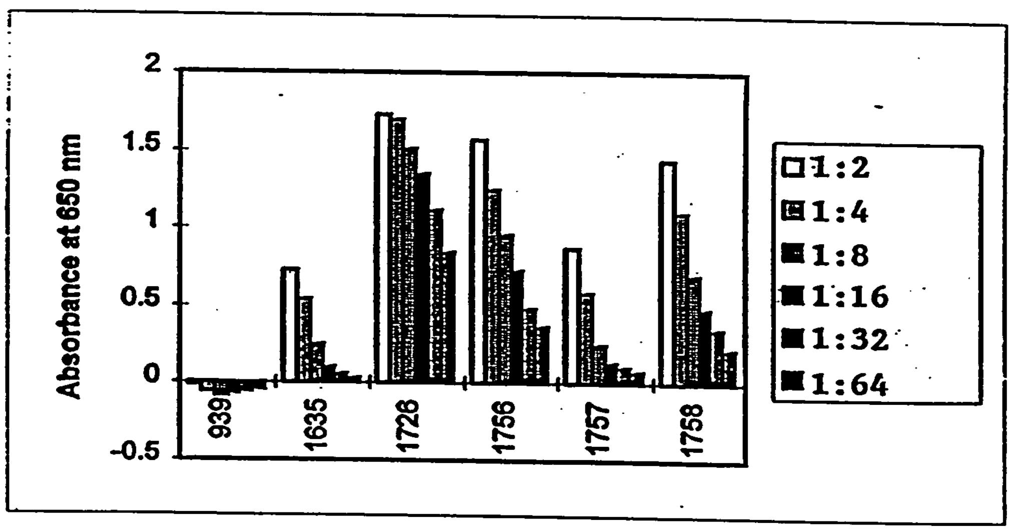

A
1
0.8
0.6
0.4
0.2
0
O.D. 650 nm
zELAN016 - hSI
zELAN016 - BSA
0.01
0.1
1
Peptide concentration ug/ml
B
0.6
0.5
0.4
0.3
0.2
0.1
0
O.D. 650 n
z024(P31)
z024 Insulin PLGA
z018(PAX2)
z018 Insulin PLGA
z021(HAX42)
z021 Insulin PLGA
z016a(SNi10)
0.01
0.1
1
Peptide concentration ug/ml

Fig. 14

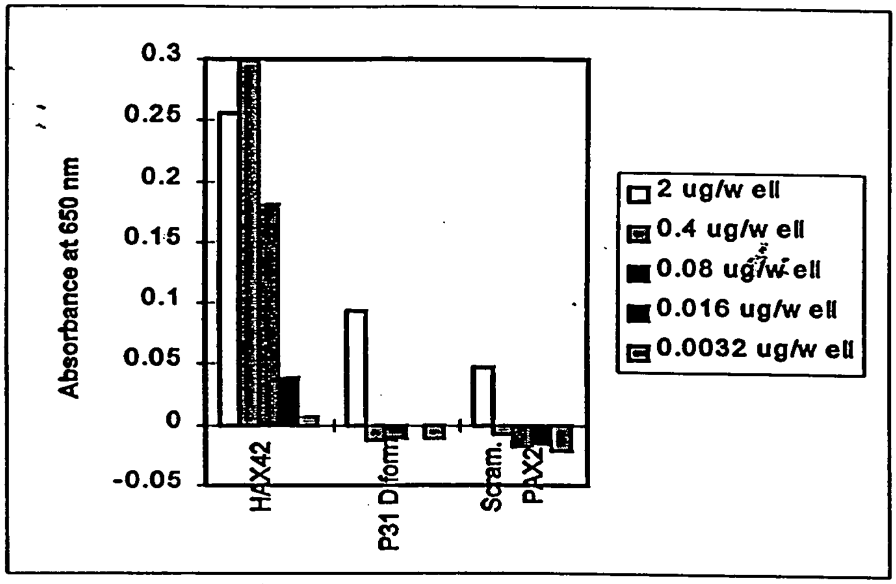

2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
Absorbance at 650 nm
939
1635
1726
1756
1757
1758
1:2
1:4
1:8
1:16
1:32
1:64

A

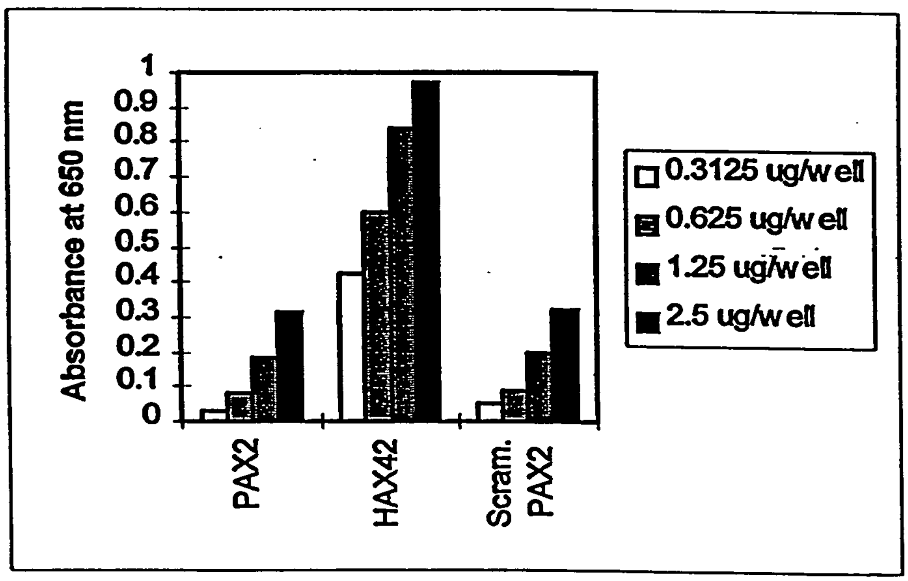

2
1.5
1
0.5
0
-0.5
Absorbance at 650 nm
939
1635
1726
1756
1757
1758
1:2
1:4
1:8
1:16
1:32
1:64

B

Fig. 15

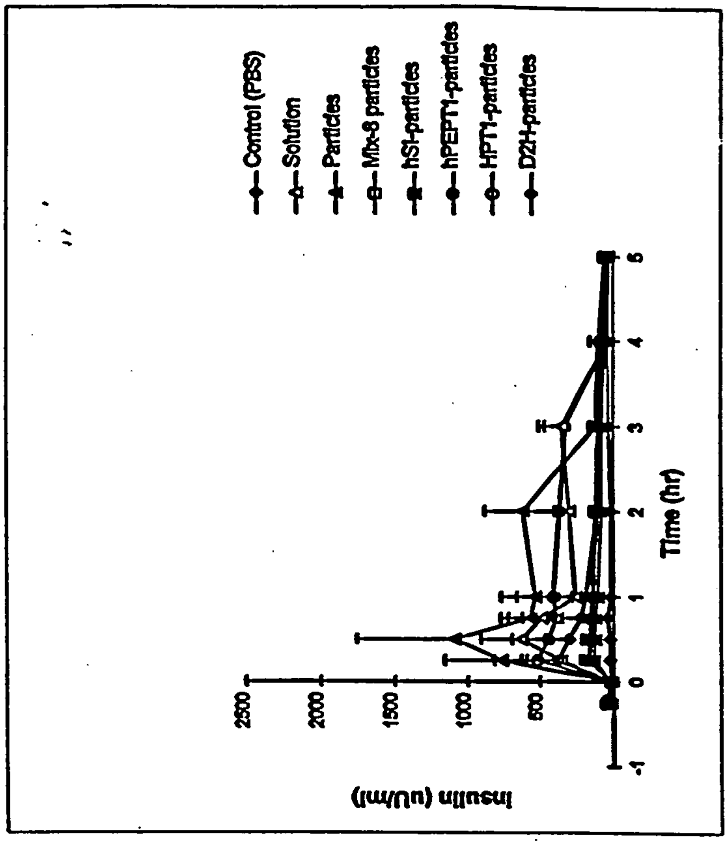

0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
Absorbance at 650 nm
HAX42
P31 Dfom
Scram. PAX2
2 ug/w ell
0.4 ug/w ell
0.08 ug/w ell
0.016 ug/w ell
0.0032 ug/w ell

A

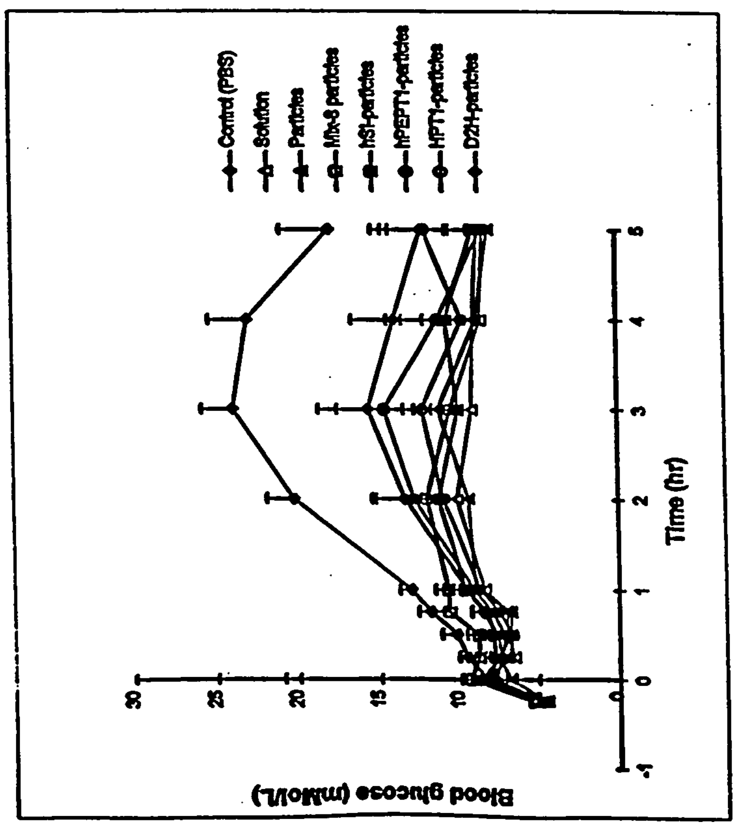

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Absorbance at 650 nm
PAX2
HAX42
Scram. PAX2
0.3125 ug/well
0.625 ug/well
1.25 ug/well
2.5 ug/well

B

Fig. 16

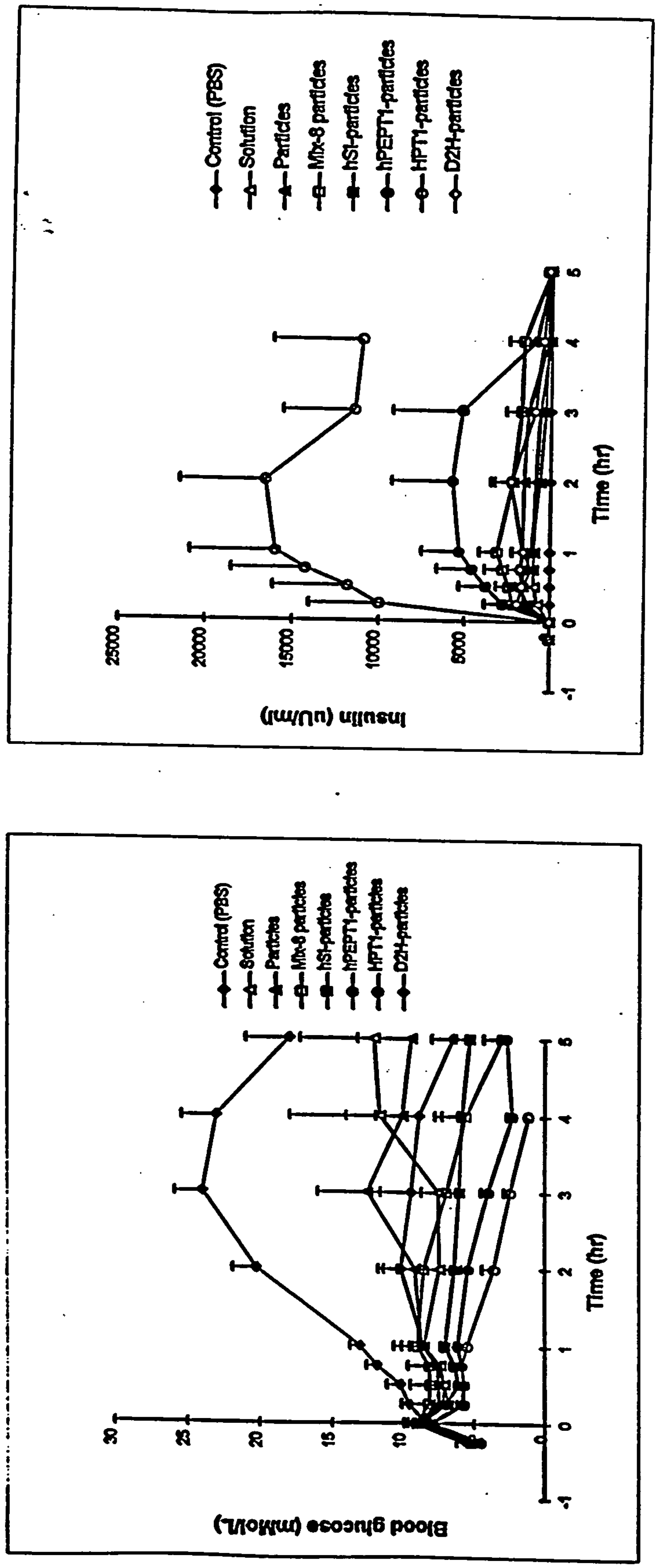

Insulin (uU/ml)
Time (hr)
Control (PBS)
Solution
Particles
Mix-8 particles
hSI-particles
hPEPT1-particles
HPT1-particles
D2H-particles

B


Blood glucose (mMol/L)
Time (hr)
Control (PBS)
Solution
Particles
Mix-8 particles
hSI-particles
hPEPT1-particles
HPT1-particles
D2H-particles

A

Fig. 17


Control (PBS)
Solution
Particles
Mix-8 particles
hSI-particles
hPEPT1-particles
HPT1-particles
D2H-particles
Insulin (uU/ml)
Time (hr)
Blood glucose (mMol/L)

A

B

Fig. 18

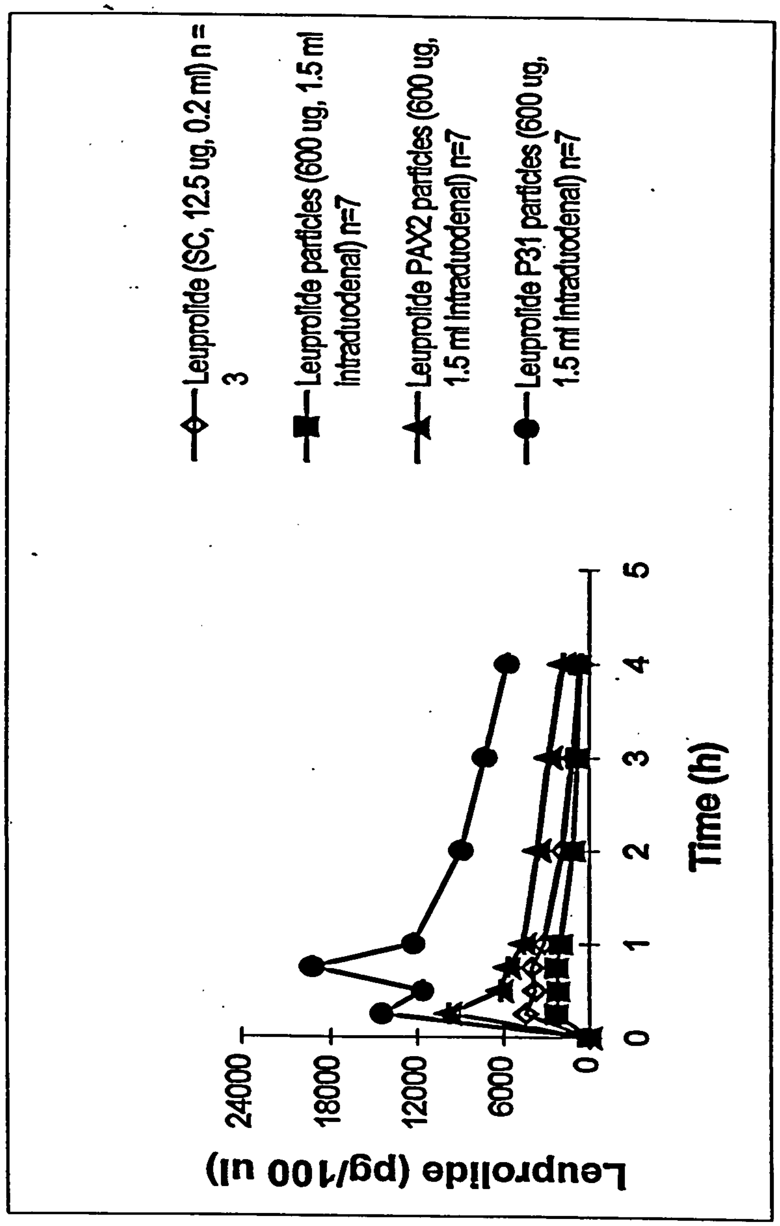

Leuprolide (pg/100 ul)
24000
18000
12000
6000
0
0
1
2
3
4
5
Time (h)
Leuprolide (SC, 12.5 ug, 0.2 ml) n = 3
Leuprolide particles (600 ug, 1.5 ml Intraduodenal) n=7
Leuprolide PAX2 particles (600 ug, 1.5 ml Intraduodenal) n=7
Leuprolide P31 particles (600 ug, 1.5 ml Intraduodenal) n=7

Fig. 19

P31 AA Seq. Position	Known Protein	Homologous Seq. Position
12-34	Fasciculin 2	10-32
4-12	Mesentericopeptidase	54-62
15-31		175-191
26-39	Core protein (Hepatitis C virus)	5-18
26-39		11-24
26-39		21-34
26-39		38-51
23-30		39-55
25-39		41-55
26-39		51-64
16-39	PT-NANBH Polyprotein N-terminus	51-64
28-40	AL2 protein (Caenorhabditiselegans)	70-82
26-38	Capsid protein (Hepatitis C virus Type 3g)	48-60
26-39	Genome polyprotein (Hepatitis C virus)	57-70

Fig. 20

DCX8AA Seq. Position	Known Protein	Homologous Seq. Position
20-27	Endo-1,4-Beta-D-Glucanase	78-85
30-37		221-228
21-34	P-Hydroxybenzoate Hydroxylase	285-298
5-15		54-64
7-21	Cytochrome	50-64
7-21	Cytochrome C3	50-64
	Trimethylarnine Dehydrogenase	208-219
32-43		396-407
30-37	Gag-JunD fusion protein	24-31
26-30		16-20
23-44	Secretin precursor, N-prosecretin, secretin ainide	18-39
33-44	T-cell receptor V beta chain	15-26
27-33		3-9
23-44	Secretin precursor pir	18-39
31-44	Hypothetical protein V (Synechocystis)	275-288
24-30		251-257
23-43	Putative RNA binding protein	230-250
28-40	Mu son of sevenless 1	1-13
24-35	Neuropeptide precursor	80-91
29-43		5-19
23-43	RNA-binding protein (Macacafascicularis)	230-250
23-43	RNA-binding protein (Homosapiens)	230-250
23-43	Autosomal gene – azoospermia factor	230-250
25-38	Collagen	25-28
24-35		4-15
29-41	Probable cell growth regulator	306-318
24-35	Ribosomal protein S2	24-35
T6-39		182-185
24-44	Caenorhabditis elegans	296-316
23-34	pid:e208155 (Homo sapiens)	61-72
36-43		116-123

Fig. 21A

DCX8A Seq. Position	Known Protein	Homologous Seq. Position
24-38	Xylulose Kinase	16-30
24-39	Caemorhabditis elegans	57-72
26-42		65-81
27-33	Hypothetical protein – phage BZ13	22-28
35-39		31-35
30-42	Cerebelllin-like glycoprotein	2-14
8-22	DNA Primase	170-184
2-7		76-81
5-21	Coat Protein (Bean common mosaic virus)	12-28
5-21	Coat protein (Bean common mossaic virus)	33-49
5-21		19-35
5-21	Polyprotein (Bean common mossaic virus)	215-231
5-21		39-55
5-21	Nib proteinlcoat protein (Cowpea aphid-bome mosaic virus)	92-108
2-13	MHC class 1 Pipi (Pithecia)	111-122
14-22		326-334
3-19	Talin (Caenorhabditis elegans)	1538-1554
2-9	Acetamidase pir	359-366
9-20		483-494
10-16	Rhizobions etli strain	134-140
17-30		173-186
31-39		200-208
2-11	Neurotoxin 1 (toxin B) A. Stokesi	7-16
12-33		26-47
21-27	Suid herpes virus 1 early protein	425-432
30-43		51-64
13-42	Rice cDNA partial sequence	50-151
8-15	Fusion protein	24-31
4-8		16-20
1-22	Secretin precursor, N-prosecretin, secretin-amide	18-39
11-22	T-cell receptor V beta chain	15-26
5-11		3-9
9-22	Hypothetical protein	275—288
2-8		251-257

Fig. 21B

DCX8A Seq. Position	Known Protein	Homologous Seq. Position
1-21	Putative RNA binding protein	230-250
6-18	Hypothetical protein-mouse pir	1-13
2-13	Neuropeptide precursor	80-91
7-21	orf3-human	5-19
1-21	RNA-binding protein	230-250
13-16	Collagen	25-28
7-19	Probable cell growth or differentiation regulator	306-318
2-13	Ribosoaml protein S2	14-25
14-17		182-185
2-22	Caenorhabditis elegans	296-316
1-12	Homosapiens	61-72
14-21		116-123
2-16	Xylulose Kinase	16-30
8-15	T cell receptor delta chain	55-62
5-8		12-15
8-17	Seq. 43 from patent US	12-21

Fig. 21C

DAB10 AA Seq. Position	Known Protein	Homologous Seq. Position
13-34	1,3-Beta-Gllucanase	231-252
3-11	Photosynthetic Reaction Center	20-28
16-27		128-139
28-35	MYB Proto-Oncogene Protein	131-138
5-18		32-45
23-36	Lysozyme Mutant	130-143
28-35	Lipase	400-407
3-15		159-171
3-37	Trypsin	169-203
13-34	1,3-1,4-Beta-Glucanase	232-253
4-10	Lactate Dehydrogenase	190-196
11-7		244-250
4-10	Apo-Lactate Dehydrogenase	190-196
11-17		244-250
4-10	Lactate Dehydrogenase	191-197
11-17		245-251
16-26	Ovotransferrin	240-250
23-36	Genome Polyprotein Matrix Protein	1022-1035
14-20	Rous sarcoma virus	43-49
2-12		13-23
14-20	Hypothetical protein-avian leukosis virus	43-49
4-20	T cell receptor delta chain variable region	1-4
14-18		12-16
2-12	Gag Polyprotein-avian endogenous virus RAV-0	139-149
14-20		169-175
	p19 Protein-avian erythroblastosis virus	189-199
14-20		219-225
7-19	ALI protein-potato yellow mosaic virus	222-234
3-22	Endo-1,4-beta glucanase	186-205
6-18	I a protein-brome mosaic virus	430-442
2-12	Gag polyprotein-Fujinami sarcoma virus	186-196
14-22		216-222
2-12	Gag protein-Rous sarcoma virus	190-200
14-20		220-226
1-12	Corticotropin-like intermediate lobe peptide	7-18
1-22	Gene product (Caenorhabditis elegans)	4-25
31-37	T cell receptor delta chain	56-62
26-39		12-15
26-37	Lysozyme Mutant	133-144

Fig. 22

```
ATG TCC CCT ATA CTA GGT TAT TGG AAA ATT AAG GGC CTT GTG CAA CCC    48
Met Ser Pro Il  Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro
1               5                   10                  15

ACT CGA CTT CTT TTG GAA TAT CTT GAA GAA AAA TAT GAA GAG CAT TTG    96
Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu
            20                  25                  30

TAT GAG CGC GAT GAA GGT GAT AAA TGG CGA AAC AAA AAG TTT GAA TTG   144
Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu
        35                  40                  45

GGT TTG GAG TTT CCC AAT CTT CCT TAT TAT ATT GAT GGT GAT GTT AAA   192
Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys
    50                  55                  60

TTA ACA CAG TCT ATG GCC ATC ATA CGT TAT ATA GCT GAC AAG CAC AAC   240
Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn
65                  70                  75                  80

ATG TTG GGT GGT TGT CCA AAA GAG CGT GCA GAG ATT TCA ATG CTT GAA   288
Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu
                85                  90                  95

GGA GCG GTT TTG GAT ATT AGA TAC GGT GTT TCG AGA ATT GCA TAT AGT   336
Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser
            100                 105                 110

AAA GAC TTT GAA ACT CTC AAA GTT GAT TTT CTT AGC AAG CTA CCT GAA   384
Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu
        115                 120                 125

ATG CTG AAA ATG TTC GAA GAT CGT TTA TGT CAT AAA ACA TAT TTA AAT   432
Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn
    130                 135                 140

GGT GAT CAT GTA ACC CAT CCT GAC TTC ATG TTG TAT GAC GCT CTT GAT   480
Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp
145                 150                 155                 160

GTT GTT TTA TAC ATG GAC CCA ATG TGC CTG GAT GCG TTC CCA AAA TTA   528
Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu
                165                 170                 175

GTT TGT TTT AAA AAA CGT ATT GAA GCT ATC CCA CAA ATT GAT AAG TAC   576
Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr
            180                 185                 190

TTG AAA TCC AGC AAG TAT ATA GCA TGG CCT TTG CAG GGC TGG CAA GCC   624
Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala
        195                 200                 205

ACG TTT GGT GGT GGC GAC CAT CCT CCA AAA TCG GAT CTG GTT CCG CGT   672
Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Val Pro Arg
    210                 215                 220

GGA TCC CCA GGA ATT CCC GGG TCG ACT CGA GCG GCC GCA TCG TGA       717
Gly Ser Pro Gly Ile Pro Gly Ser Thr Arg Ala Ala Ala Ser
225                 230                 235
```

Fig. 23